

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2001, 09:31:40 ; Search time 1961.35 Seconds
(Without alignments)
11553.396 Million cell updates/sec

Title: us-09-484-964-1

Perfect score: 1465
Sequence: 1 cgggaagagttgttaaccc.....gataatgcatgcagcc 1465

Scoring table: OLIGO-NUC
Gapop 60.0 , Gapext 60.0

Searched: 1344157 seqs, 7733874588 residues

Word size : 50

Total number of hits satisfying chosen parameters: 29

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_ov:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_p11:*
13: gb_p12:*
14: gb_p13:*
15: gb_p14:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_hugo_hum:*
20: em_hugo_inv:*
21: em_hugo_rod:*
22: em_hugo_hum1:*
23: em_hug_hum2:*
24: em_hug_hum3:*
25: em_hug_hum4:*
26: em_hug_hum5:*
27: em_hug_hum6:*
28: em_hug_hum7:*
29: em_hug_hum8:*
30: em_hug_inv1:*
31: em_hug_inv2:*
32: em_hug_other:*
33: em_hug_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_ov:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_v1:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_v11:*
59: gb_v12:*
60: gb_hug1:*
61: gb_hug2:*
62: gb_hug3:*
63: gb_hug4:*
64: gb_hug5:*
65: gb_hug6:*
66: gb_hug7:*
67: gb_hug8:*
68: gb_hug9:*
69: gb_hug10:*
70: gb_hug11:*
71: gb_hug12:*
72: gb_hug13:*
73: gb_hug14:*
74: gb_hug15:*
75: gb_hug16:*
76: gb_hug17:*
77: gb_hug18:*
78: gb_hug19:*
79: gb_hug20:*
80: gb_hug21:*
81: gb_hug22:*
82: gb_hug23:*
83: gb_hug24:*
84: gb_hug25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_r01:*
95: gb_r02:*
96: gb_in4:*
97: gb_pi10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778	53.1	1017	97	HSU38784
2	775	52.9	1223	97	HSU61397
3	775	52.9	1227	91	BC006462
4	679	46.3	816	97	HSU67132
5	625	42.7	160939	76	AC079354
6	620	42.3	1514	9	AR087051
7	620	42.3	1514	97	HSU83117
8	343	23.4	590	93	HSSMT3C
					X99586 H.sapiens m

9	284	19.4	372	9	AX021844
10	255	17.4	306	97	HSU72722
11	150	10.2	155913	81	AL513282
12	150	10.2	167440	80	AL359981
13	98	6.7	106168	87	AC011450
14	98	6.7	178733	73	AC068786
15	97	6.6	181861	72	AC064836
16	65	4.4	510	7	AF242526
17	65	4.4	158519	89	AL135926
18	62	4.2	1187	94	AF033353
19	62	4.2	180821	68	AC023974
20	62	4.2	205741	60	AC008548
21	62	4.2	235411	76	AC084066
22	58	4.0	124764	60	AC008377
23	58	4.0	146267	60	AC008815
24	58	4.0	160970	61	AC010501
25	58	4.0	161190	64	AC016567
26	58	4.0	163082	67	AC022142
27	58	4.0	173048	87	AC008653
28	58	4.0	209022	71	AC034212
29	52	3.5	162776	66	AC021171

ALIGNMENTS

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	ALIGNMENTS
1	HSU38784	1017 bp	MRNA	PRI	16-OCT-1996			
1	HSU38784	Human ubiquitin-like protein mRNA, complete cds.						

U38784.1 GI:1574947

human.

Human sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Shen, Z., Pardington-Purtymun, P.E., Comeaux, J.C., Moyzis, R.K. and Chen, D.J.

UBLI, a human ubiquitin-like protein associating with human RAD51/RAD52 proteins

JOURNAL Genomics, 36 (2), 271-279 (1996)

96411684

2 (bases 1 to 1017)

Direct Submission

Submitted (16-OCT-1995) Zhiyuan Shen, Life Sciences Division, Los Alamos National Lab, MS M888, Los Alamos, NM 87545, USA

FEATURES

source

1. 1017

/organism="Homo sapiens"

/db_xref="taxon:9606"

/note="This cDNA was cloned with the yeast two hybrid system. The protein interacts with human RAD51 and RAD52 proteins in yeast"

67. 372

/note="similar to ubiquitin and to yeast Smtp (suppressor of Mif2); Method: conceptual translation supplied by author"

/codon_start=1

/protein_id="AAC50733.1"

/db_xref="GI:1574948"

/translation="MSDQKAPKSTEDLDGKKEGYIKLVIGQSSSEHFVKKMTTLKKIKSEYCOQGVPMNSRFLRFGORIDNHPKELGMEEDVLEIVQETGSHSTV"

BASE COUNT 310 a 180 c 213 g 314 t

ORIGIN

Query Match 53.1% Score 778; DB 97; Length 1017;

Best Local Similarity (99.9%) Pred. No. 0;

Matches 898; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY	140	aagtggaatattataaactcaagtcattgacagatagcagtgatctcaag	199
Db	119	AAGTGGAATATTATTAACTCAAGTCATTGGACAGATGAGTGAGTACTTCAAG	178
QY	200	tgaatagacaacacatctcaagaactcaagaatcaacgctcaagaagcgtgttc	259
Db	179	TGAATAAGACAACATCTCAAGAACTCAAGAAATCAATACGTCAAGACGGGTTC	238
QY	260	caatgaattacacacagttctcttgaaggtcagaagaattgcgtaatactcaaa	319
Db	239	CAATGAATTAATCTCAAGGTCTCTTGAAGGTCAGAGAACTCTATATATATCCAA	298
QY	320	aagaactgggaatggagaagaagatgattgaagtatacagaacaacgggggtc	379
Db	299	AAGAAGTGGGAATGGAGGAAGAAGATGATGAACTTATAGGAACAAACGGGGGTC	358
QY	380	atcaacagtttagatattcttttatttttttttttttttttttttttttttttt	439
Db	359	ATTCAACAGTTTAGATATTTCTTTTATTTTCTTTTCTTTTCTTTTCTTTTCTTT	418
QY	440	taaaaatgcttcttcttgaatgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	499
Db	419	TAAATATGTTCTTTTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	478
QY	500	cttgaacacatctgttaattgaattcagtgctcattatcattatgtgtgtttcat	559
Db	479	CTTGAACATCTGTGAATTTGAATTTGATTTGATTTGATTTGATTTGATTTGATTTG	538
QY	560	tgtcagatcttctgtgacacagctcagtcctcctcattatcctcctcctttaa	619
Db	539	TGTGCTGATTTTGTGTGATCAAGCTCAGTCCCTTCAATATACCTCTCTTTTAAA	598
QY	620	attaagtgacacagagagtgacacacacacacacacacacacacacacacacac	679
Db	599	ATTAGGTGTGACACAGAGAGTGACCTTTTTCAGACATTTGATTTTTCAGCTGTGTG	658
QY	680	taataagatcagcaacagagagtgacacacacacacacacacacacacacacac	738
Db	659	TAAATATGATGACCAATGCAAGGTGTGATATGATATGATATGATATGATATGATATG	718
QY	739	gcattgattcattcactcctcgtgacacacacacacacacacacacacacacac	798
Db	719	GCATGTGATTTCTTCACTCTGACCTGACCTGACCTGACCTGACCTGACCTGACCTG	778
QY	799	aactgaactgtggaac	858
Db	779	AAGTGAATGTGGAACAAATGACCTTCTTCACTGAAAGCTTTTAAATTTTGAAGGT	838
QY	859	ctggaac	918
Db	839	CTGGAACCAAGAGAGAGATATCAGGTGAACTCAAGATGACAGATATGAGTGAAGTGA	898
QY	919	tgactaactccaagaatgagcttcaacgagaagaagcatttaagaatttaaaatct	978
Db	899	TGACTAACTCCAAAGAGAGGTCTTCACTGAAAGGCAATTTTAAAGATTTTAAAAA	958
QY	979	tgtcagaagaatcccaagaagaatcctaatttcaatgaacatttaaaacatcaatc	1037
Db	959	TGTGAGAAGATCCCAAGAAAGTCTTAATTTTCAATTGCAATTTAATTAAGCTATGATCC	1017

RESULT 2

LOCUS HSU61397 1223 bp

DEFINITION Human ubiquitin-homology domain protein P101 mRNA, complete cds.

ACCESSION U61397

VERSION U61397.1

KEYWORDS

SOURCE

ORGANISM

human.

Human sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 1223)

AUTHORS Boddly, M.N., Howe, K., Etkin, L.D., Solomon, E. and Freemont, P.S.
 TITLE PIC 1, a novel ubiquitin-like protein which interacts with the pML
 component of a multiprotein complex that is disrupted in acute
 promyelocytic leukaemia
 JOURNAL Oncogene 13 (5), 971-982 (1996)
 MEDLINE 96400311
 REFERENCE 2 (bases 1 to 1223)
 AUTHORS Howe, K., Boddly, M.N., Etkin, L.D., Solomon, E. and Freemont, P.S.
 TITLE Direct Submission
 JOURNAL Submitted (20-JUN-1996) Somatic Cell Genetics, Imperial Cancer
 Research Fund, 44 Lincoln's Inn Fields, London WC2A 3PX, UK
 FEATURES
 source Location/Qualifiers
 1..1223
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /issue_type="placenta"
 146..451
 /codon_start=1
 /product="ubiquitin-homology domain protein PIC1"
 /protein_id="AAB40388.1"
 /db_xref="GI:1518694"
 /translation="MSDQEAAPSTEDIGDKKEEYIKLYIGDSSPIHPKVMTHL
 KLNESICQROGVPMNSLRFLESGRIADNHTRELGMEEEDYIEVYQEDTGSHTV"
 BASE COUNT 371 a 209 c 268 g 375 t
 ORIGIN

Query Match 52.9%; Score 775; DB 97; Length 1223;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1065; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

```

4 gaagattgttaaaccccgagagaggttcttaccgagcgctgtgtgcgaga 63
Db 62 GAAGATTGTTAACCCCGAGAGAGGTCTCTACCGAGCGCGTGTGCGGAGA 121
64 ccccggtgtgaagcaccgcatcatcgtctgacccgagagcaaaccttaagac 123
Db 122 CCCCcgggtgaagcaccgcatcatcgtctgacccgagagcaaaccttaagac 181
124 ttggggataagaagaaggtgaatatataaactcaagtcattgagacagatgac 183
Db 182 TTGGGGATAGAGAGAGGTGAATATATTAACCTCAAGTCTATTGGACAGATAGCAGT 241
184 gagatcaactcaagtgaaatgacaacacatctcaagaacacaaatcactactgt 243
Db 242 GAGATCACTTCAAGTGAATGACACACATCTCAAGAACTCAAGATCAATCTGT 301
244 caagacagaggtgttccaatgaatcactcaggttctcttggagggtcagagatgct 303
Db 302 CAAAGACAGGGGTTCCTCAATGAATTCACCTCAGGTTCCTTTGAGGGTCAGAGATTCGT 361
304 gataataactcaaaagaagacgtggaatgagagaagatgtgattgaattatcag 363
Db 362 GATATATCTACTCCAAAGAACTGGGAATGAGAGAAAGATGTGATGAAGTTATCAG 421
364 gaacaacaggggtgtcattcaacagtttagatatcttcttattcttcttccctc 423
Db 422 GAACAACAGGGGTGTCAATCAAGTTAGATTCCTTTTATTTTTCCTTCCCTCC 481
424 aatccttttttttttttttttttttttttttttttttttttttttttttttt 483
Db 482 AATCCTTTTATTTTAAATATGTTCTTTTGAATGTGTCTTCAAAAGGAAATTTGAA 541
484 aactgacacccatctcttgaacaacatgttaattgaattcctagtgctcattatcat 543
Db 542 AACTGACACCCATCTCTTGAACATCTGTGAATTTGAATTCAGTCTCATTTATTCAT 601
544 taattgttttttttttttttttttttttttttttttttttttttttttttttt 603
Db 602 TATTGTGTTTTCATTTGCTGATTTTGGTGAATCAAGCTCAGTCCCTTATATTTAC 661
604 cctcccttttaaaatagctgtgacagagagtcacatttcaagacattgacatt 663

```

```

Db 662 CCTCTCTTTTAAAAATTTACGTGTGCACAGAGAGTCCACTTTTTCAGGACATTCATT 721
664 ttcaagctgtgtgtataaataaagaatgacccaatgacaaagtgttcaatgaacttccaat 723
Db 722 TTCAGCGCTGTGTGATTAATTAATGATGACCAATGCAAGGTTCATTAATCACTTCCAA 781
724 tggccctgatgttc-agaatgtgtattcttaccctc-aggacgtgtgacttgaaggag 781
Db 782 TGGCCCTGATGTTCAGATGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 841
782 atggaagttttcagagacagcagactgtgagaaaatgacattccttacttgaagctac 841
Db 842 ATGGAAGTTTTCAGAGAACTGAACTGTGAAAATGACCTTCTTAACCTGAAGCTAC 901
842 tttaaaatttgaagggtctgagccaaaagaagaaggaatataggttgaagtaagatgac 901
Db 902 TTTTAAATTTTGAAGGTCTGACCCAAAAGAGAGAAATATAGTTGATGAAGCAAGATGAC 961
902 agataagtgagaagatgactgaactccaagaatgaggtctcactgaagaagaagcattta 961
Db 962 AGATAAGTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1021
962 agatttttaaaatctgttcagagaagatcccaagaagaagttcatttcatltagcaatta 1021
Db 1022 AGATTTTAAAAAATCTGTGAGAGATCCCGAAGAAAGTTCTAATTTTCATTAACCAATTA 1081
1022 ataagctatactcagagaatgataacacagagaacactgctcttltt 1069
Db 1082 ATAACTTATACATGACGAAATGAATACAAAGAACCTGCTCTTTT 1129

```

RESULT 3
 BC006462 1227 bp mRNA PRI 13-APR-2001
 LOCUS Homo sapiens, ubiquitin-like 1 (sentrin), Clone MGC:2095, mRNA,
 DEFINITION complete cds.
 ACCESSION BC006462
 VERSION BC006462.1 GI:13623670
 KEYWORDS MGC.
 SOURCE human.
 ORGANISM Homo sapiens
 Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1 (bases 1 to 1227)
 REFERENCE Strausberg, R.
 Direct Submission
 Submitted (09-APR-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 CONTACT: Robert Strausberg, Ph.D.
 Tel: (301) 496-1550
 Email: Robert.Strausberg@nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
 DNA Sequencing by: Genome Sequence Centre,
 BC Cancer Agency, Vancouver, BC, Canada
 Info@egsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
 Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
 Leticia Hsiao, Martin Krzyzinski, Reta Kutsche, Oliver Lee, Soo
 Sen Lee, Victor Ling, Carrie Mathewson, Candice McEavey, Steven
 Ness, Pawan Pandoh, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
 Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stolt,
 Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
 George Yang, Scott Zuyderduyn, Marco Marra.

REMARK
 COMMENT

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LINT at: <http://image.llnl.gov>
 Series: IRAL Plate: 8 Row: f Column: 12.
 Location/Qualifiers

source

1. 1227

/organism="Homo sapiens"

/db_xref="taxon:9606"

/db_xref="taxon:9606"

/clone="MGC:2095 IMAGE:3546125"

/tissue_type="Placenta, Chorioncarcinoma"

/clone_id="MGC_21"

/lab_host="DH10B-R"

/note="Vector: pORF7"

132. 437

/product="ubiquitin-like 1 (sentrin)"

/protein_id="AAH06462.1"

/db_xref="GI:13623671"

/translation="MSDQKASTEDLGKKKEEYIKLVIGDSSSEIHFVKMTTHL
KKIKSYCORGVPNMSLRFLREGORLADNHTPELMGEEDVIEVYEOGTGHSIV"

BASE COUNT 386 a 208 c 259 g 374 t

ORIGIN

Query Match 52.9%; Score 775; DB 91; Length 1227;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1065; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 4 gaagagattgtaaaccccgagcgaggtctgtcttaaccgagcgcgctgtgtgcgaga 63
DB 48 GAAGGATTGTGTAACCCCGAGCGAGGTTCTGCTTACCCGAGCGCGCTGTGTGCGAGA 107
QY 64 ccccgaggtgaagccacccgtcatcatgtctgaccagagagcaaacctcaactgagac 123
DB 108 CCCCAGGTGAAGCCACCGCTCATCATGTCTGACACAGAGCAAAACCTTCAACTGAGAC 167
QY 124 ttggggataaagaagaagtgatataataactcaagctatgtgacagatagcagt 183
DB 168 TTGGGGATTAAGAAGAAGTGATATTAATCAATCAAGCTATGACAGATAGCAGT 227
QY 184 ggaattcactcaagaatgataacacacatctcaagaatacaagaatcactgt 243
DB 228 GAGATTCACTTAAGTGAATGACACACATCTCAAGAACTCAAGAACTCAATACAGT 287
QY 244 caaagacagaggtgtcccaatgaatcactcaagttctcttggaggtcagaagaattc 303
DB 288 CAAAGACAGAGGTGTCTCAATGATTCATCAAGTTCTCTTGAAGGTCAGGAATTCCT 347
QY 304 gataatcactcaaaagaactggagatgagagagaagatgtgataatgataag 363
DB 348 GATTAATCATATCCAAAGAAGTGGGAATGAGGAAGAAGATGATGATGATGATGAT 407
QY 364 gacaaagaggggtgcatcaacagtttagatatctcttta-ttttttcttccct 422
DB 408 GAACAACGGGGGTCAATCAACAGTTAGATATCTTTTATTTTCTTTCCCT 467
QY 423 caatccttttttttaaaatagttcttttgaatgtgtgttcaaacgaaatgta 482
DB 468 CAATCCTTTTATTTTAAATAGTTCTTTGTAATGTGTGTGTAACGGAATGTA 527
QY 483 aaactggaacccatctcttgaacacatctgtaattgtaattcagtgctcatatca 542
DB 528 AAACAGGACCCCATCTCTTGAAGAACATCTGTATTTGAATGCTAGGCTCATTTTCA 587
QY 543 ttattgtgttttcaatgtgtcgaattttgtgtatcaagcctagtcctcatatla 602
DB 588 TTATTGTGTTTCAATGTGCTGATTTTGTGATCAAGCCTCAGTCCCTCATATTTA 647
QY 603 cccctccttttaaaatcgtgtcaacagagagtccttttcaagaaatgcat 662
DB 648 CCCCTCTCTTTTAAATATTCGTGTGACAGAGAGGTTCACCTTTTACAGACATTGCAT 707
QY 663 tttaagctgtgtgtgataaataagatcgaacaaatgtaagtgatcatatgacttccaa 722
DB 708 TTTCAGGCTGTGTGTAATATAGATGACACATGCAAGTGTTCATTAATGACTTTTCAA 767
QY 723 ttggcctgtagtgc-agaatgtgataactcaactcctgtagctgtgacttcaagtgagag 781

DB 768 TTGGCCCTGATGTTTACGATGATTAATCTTCACTTCGAGTGTGACTTTCAGTGGAG 827
QY 782 atggaagtttccaagaactgaactgtggaataatgacttcttcaactgaagctac 841
DB 828 ATGGAAGTTTTCAGAGAACTGAGTGTGGAATAATGACCTTCTTAACTTGAAGCTAC 887
QY 842 tttaaaattgaaggtctgagccaaagaagaagaatcaggttgaagcagaatgac 901
DB 888 TTTTAAATTTGAGGCTGTGACCAAAAGAGCAATATCAGTTGATGATCAAGATGAC 947
QY 902 agataagtgagatgaatgaactactccaagaatggtcttcaactgaagaagacatlla 961
DB 948 AGATAGGAGAGATGATGACTAATCCCAAGATGCTTCACTGAAGAAAGGCAATTTTA 1007
QY 962 agatttttaaaactgtgtcgaagaatcccaagaagaatttcaatttcaatgaatla 1021
DB 1008 AGATTTTAAATCTTGTGTCAGAGATCCAGAAAGTTCTTAATTTCAATTAACCAATTA 1067
QY 1022 ataaagctatacatgcaagaatgaataacacagaacactgctcttlt 1069
DB 1068 ATAAAGCTATACATGCAAGAAATGATATACACAGAAACACTGCTCTTTT 1115

RESULT 4
LOCUS HS067122 816 bp mRNA PRI 15-FEB-1997
DEFINITION Human ubiquitin-related protein SUMO-1 mRNA, complete cds.
VERSION U67122.1 GI:1762972
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 816)
Mahajan R., Delphin C., Guan T., Gerace L. and Melchior F.
A small ubiquitin-related polypeptide involved in targeting RanGAP1
to nuclear pore complex protein RanBP2
Cell 88 (1), 97-107 (1997)
JOURNAL MEDLINE
2 (bases 1 to 816)
Mahajan R., Guan T., Delphin C., Gerace L. and Melchior F.
Direct Submission
Submitted (16-AUG-1996) Department of Cell Biology, The Scripps
Research Institute, 10666 N. Torrey Pines Road, La Jolla, CA 92037,
USA

FEATURES
source
Location/Qualifiers
1..816
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE Consortium Clone ID: 49768, 51818, 199507"

CDS

81..386
/note="conjugated post-translationally to RanGAP1;
ubiquitin-related protein; similar to UBL1 encoded by
GenBank Accession Number U36784, P1C1 encoded by GenBank
Accession Number U61397 and GmP1 encoded by GenBank
Accession Number U72722"
/product="SUMO-1"
/protein_id="AAC5096.1"
/db_xref="GI:1762973"
/translation="MSDQKASTEDLGKKKEEYIKLVIGDSSSEIHFVKMTTHL
KKIKSYCORGVPNMSLRFLREGORLADNHTPELMGEEDVIEVYEOGTGHSIV"

BASE COUNT 235 a 153 c 173 g 255 t

ORIGIN

Query Match 46.3%; Score 679; DB 97; Length 816;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 729; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 8 gatitgtaaaccccgagcgaggtctgtcttaaccgagcgcgctgtgtgcgagacccc 67

OY	422	tcacatccttttattttaaanaagtagtcttcttgtaatggtgttccaagaagattg	481
Db	470	TCATATCCTTTTATTTTAAAAATAGATCTTTTGTATGTAGTGGGTCAAAACGGAATTG	529
OY	482	aaaactggaaccccatctcttgaagaactcggaattgattgattagtagtgcattatcc	541
Db	530	AAAACGTGCAACCCCACTCTTTGAAACATCTGTATTTGAATTCGTAGTCCTCATATTTC	589
OY	542	attatgttcttttcatctgctgattttgttgatcaaacctccatcccttcatatt	601
Db	590	ATTATTTGTTTGTTCATGTGCTGATTTTGTGATCATCAACCTCAGTCCCTTCATATT	649
OY	602	accctctcccttttaaaattagctgctgacagagaggctcaaccttttcaggacattgca	661
Db	650	ACCCCTCCCTTTTAAAAATTACGTGTGACAGAGAGGTGACACCTTTTACAGACATTGCA	709
OY	662	ttttcaagcctt 672	
Db	710	TTTTTCAGGCTT 720	
RESULT	8		
LOCUS	HSMT3C	590 bp	19-MAR-1997
DEFINITION	H.sapiens mRNA for SMT3C protein.		
ACCESSION	X99586		
VERSION	X99586.1 GI:1770520		
KEYWORDS	SMT3C gene; suppressor; ubiquitin-like protein.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	1 (bases 1 to 590)		
TITLE	Lapenta, V., Chiurazzi, P., van der Spek, P., Pizzuti, A., Hanaoka, F. and Brhe, C.		
JOURNAL	SMT3A, a human homologue of the S. cerevisiae SMT3 gene, maps to chromosome 21qter and defines a novel gene family		
MEDLINE	Genomics 40 (2), 362-366 (1997)		
REFERENCE	2 (bases 1 to 590)		
AUTHORS	Chiurazzi, P.		
JOURNAL	Direct Submission		
FEATURES	Submitted (26-JUL-1996) P. Chiurazzi, Universita' Cattolica - Roma, largo F. Vito 1, I- 00168 Roma, ITALY		
SOURCE	location/Qualifiers		
gene	1. 590		
CDS	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/dev_stage="Foetus"		
	/tissue_type="Brain"		
	42. 347		
	/gene="SMT3C"		
	42. 347		
	/gene="SMT3C"		
	/function="suppressor of MIF2 which encodes a centromere protein"		
	/note="ubiquitin-like protein"		
	/codon_start=1		
	/product="SMT3C protein"		
	/protein_id="CAA67898.1"		
	/db_xref="GI:1770521"		
	/db_xref="SWISS-PROT:Q93068"		
	/translation="MSNOEAKPSTEDLGKKEGEYIKLVIGDSSIEHKVKTTHL		
	KKLESTQKRGVPMNSIRLFLPGORIANHNPKEIGDEEDVLEVVOEQTGHSTV"		
BASE COUNT	168 a 113 c 126 g 170 t 13 others		
ORIGIN			
Query Match	23.4%	Score 343;	DB 93; Length 590;
Best Local Similarity	99.7%	Pred. No. 2.9e-161;	
Matches 393; Conservative	0;	Mismatches 1;	Indels 0; Gaps 0.

OY 388 gtttag 393
 |||||
 Db 301 GTTTAG 306

RESULT 11
 AL513282 155913 bp DNA HTG 26-FEB-2001
 LOCUS Homo sapiens chromosome 1 clone RP11-179G5, *** SEQUENCING IN
 DEFINITION PROGRESS ***, 8 unordered pieces.

ACCESSION AL513282.5 GI:13162044
 VERSION AL513282
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 155913)

REFERENCE 1
 Pavlitt, R.
 Direct Submission
 Submitted (25-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Feb 28, 2001 this sequence version replaced g1:13161719.

COMMENT
 ----- Genome Center
 Center: Sanger Centre
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk
 Project Information
 Center project name: bA179G5
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Sequencing vector: plasmid; L08752; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Consensus quality: 153015 bases at least Q40
 Consensus quality: 153938 bases at least Q30
 Consensus quality: 154565 bases at least Q20
 Insert size: 155213; sum-of-contigs
 Insert size: 164818; 1.0% error; agarose-fp
 Quality coverage: 6.88x in Q20 bases; sum-of-contigs Quality
 coverage: 7.27x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 8 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 20838: contig of 20838 bp in length
 20839 20938: gap of 100 bp
 20939 50589: contig of 29651 bp in length
 50590 50689: gap of 100 bp
 50690 81538: contig of 30849 bp in length
 81539 81638: gap of 100 bp
 81639 89682: contig of 8044 bp in length
 89683 89782: gap of 100 bp
 89783 108105: contig of 18323 bp in length
 108106 108205: gap of 100 bp
 108206 149395: contig of 41190 bp in length
 149396 149495: gap of 100 bp
 149496 153075: contig of 3580 bp in length
 153076 153175: gap of 100 bp
 153176 155913: contig of 2738 bp in length.

Location/Qualifiers
 1. 155913
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-179G5"

misc_feature
 1..20838
 /note="assembly_fragment:02743
 fragment_chain:1
 clone_end:17
 vector_side:left"
 20939..50589
 /note="assembly_fragment:01236
 fragment_chain:1"
 50690..81538
 /note="assembly_fragment:00478
 fragment_chain:2"
 81639..89682
 /note="assembly_fragment:03283
 fragment_chain:2"
 89783..108105
 /note="assembly_fragment:00942
 fragment_chain:3"
 108206..149395
 /note="assembly_fragment:02750
 fragment_chain:3"
 149496..153075
 /note="assembly_fragment:02486
 fragment_chain:4"
 153176..155913
 /note="assembly_fragment:03083
 fragment_chain:4"

BASE COUNT 43830 a 34529 c 35326 g 41527 t 701 others
 ORIGIN

Query Match 10.2%; Score 150; DB 81; Length 155913;
 Best Local Similarity 98.8%; Pred. No. 4.1e-64;
 Matches 400; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 27 gaggttcgtcttaaccgagcgctgtgtgagagaccggtggaagccacgctcat 86
 |||||
 Db 136416 GAGGTCGCTTACCCGAGCGCGCTGTGCGGAGACCCCGGCGGAGGCCACTGTCTAT 136357

OY 87 catgtctgaccagagagcaaaccttcaactgagagcttggggataagaagaagtga 146
 |||||
 Db 136336 CATGCTCACGACGAGGAGGAAACCTTCACTGAGGAGACTTGGGGATTAAGAGGAAGGTGA 136297

OY 147 atataatcaactcaagtcattcttgacagagatgagcagtgagagatcactcaagtgaat 206
 |||||
 Db 136236 ATATATTAAACCAAAAGCATATGACAGGATGACAGATTCATCTCAAGTGAAT 136237

OY 207 gacacacatctcaagaactcaagaatcatalactgtcaagaagcagggttccaatga 266
 |||||
 Db 136236 GACACACATCTCAAGAAACCAAAAGTCACTGTCACAAAGACAGGCGTCCAAATGAA 136177

OY 267 tttaactaggtttctcttgaggttcaagaatgttgatataactactcaaaagaact 326
 |||||
 Db 136176 TTCATTCAAGTTTCCTTTGAGGTCAGAGAAATTCGTATATCTACTCCAAAGAACT 136117

OY 327 gggaaatggagagaagatgatgtgaagttatcaagaacacggggggtcattcaac 366
 |||||
 Db 136116 GGGAAATGGAGAGAAAGATGATGAGTTATCAGAGAAATGGGGGCTCATTCAC 136057

OY 387 agttagatattctttatttttttttttttttttttttttttttttttttttttt 431
 |||||
 Db 136056 AGTTAGATATTCTTTTATTATTTTTCCTTCCTCAATCTTT 136012

RESULT 12
 AL359981 167440 bp DNA HTG 20-JAN-2001
 LOCUS Homo sapiens chromosome 1 clone RP11-115N23, *** SEQUENCING IN
 DEFINITION PROGRESS ***, 7 unordered pieces.

ACCESSION AL359981
 VERSION AL359981.10 GI:12331082
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE human.

ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 167440)
 AUTHORS Pavlic, R.
 TITLE Direct Submission
 JOURNAL Submitted (20-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
 On Jan 22, 2001 this sequence version replaced gi:10186747.
 COMMENT ----- Genome Center
 Center: Sanger Centre
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquerry@sanger.ac.uk
 Project Information
 Center project name: h115N23
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Sequencing vector: plasmid; 108752; 100% of reads
 Chemistry: Dye-terminator Big Dye; 83% of reads
 Consensus quality: 165078 bases at least Q40
 Consensus quality: 165122 bases at least Q30
 Consensus quality: 16518 bases at least Q20
 Insert size: 166840; sum-of-contrigs
 Quality coverage: 6.32x in Q20 bases; sum-of-contrigs Quality
 coverage: 8.08x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 7 contrigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contrigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 17991: contrig of 17991 bp in length
 * 17992 18091: gap of 100 bp
 * 18092 34810: contrig of 16719 bp in length
 * 34811 34910: gap of 100 bp
 * 34911 125180: contrig of 90270 bp in length
 * 125181 125280: gap of 100 bp
 * 125281 129032: contrig of 3752 bp in length
 * 129033 129132: gap of 100 bp
 * 129133 132599: contrig of 3467 bp in length
 * 132600 132699: gap of 100 bp
 * 132700 141603: contrig of 8904 bp in length
 * 141604 141703: gap of 100 bp
 * 141704 167440: contrig of 25737 bp in length.
 Location/Qualifiers
 1. 167440
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone="RP11-115N23"
 /clone_1b="RPC1-11.1"
 1. 17991
 /note="assembly_fragment:00134
 fragment_chain:1"
 18092. 34810
 /note="assembly_fragment:02116
 fragment_chain:1"
 34911. 125180
 /note="assembly_fragment:03000
 fragment_chain:1"
 125281. 129032
 /note="assembly_fragment:01240
 fragment_chain:1"
 129133. 132599
 /note="assembly_fragment:00856
 fragment_chain:1"

misc_feature 132700..141603
 /note="assembly_fragment:00528"
 misc_feature 141704..167440
 /note="assembly_fragment:02701"
 BASE COUNT 47457 a 37675 c 35802 g 45906 t 600 others
 ORIGIN
 Query Match 10.2%; Score 150; DB 80; Length 167440;
 Best Local Similarity 98.8%; Pred. No. 4.1e-64;
 Matches 400; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
 QY 27 gaagctctcttaaccgagcgctgctggtggagaccgaggtgaaccgacat 86
 Db 154488 GAGCTTCGCTGCTTACCGAGCCGCTGCTGCGAGACCCCGGGGTGAAGCCACTGTCA 154429
 QY 87 catgtctgaccgagggagcaaaccttcaactgaagactgggggataagaagcaagtga 146
 Db 154428 CATGCTGACGAGGAGGCAAACTTCAACTGAGGACTTGGGGATAGAAGAGAGGTGA 154369
 QY 147 atataataacccaagatcattggagcagatagaagagagattcaactcaagtgaat 206
 Db 154368 ATATATTAACCTCAAGATCTTGGACAGGATAGCAGTGAATTCCTCAAGTGAAT 154309
 QY 207 gcaacacatctcaagaactcaagaatcatatctgcaagaagaggtgttccaatgaa 266
 Db 154308 GACAAACATCTCAAGAACTCAAAAGATCATCTGTCAAAGACAGGCGTTCCTCAATGAA 154249
 QY 267 ttcactcaagttctctcttgggtgagagatgctgataatcaactcaagaact 326
 Db 154248 TTCATTACAGGTTCTCTCTTGGGTGAGAGATGCTGATTAATCAATCAAAAGACT 154189
 QY 327 gggaaatgaggaagaatgtagtgaattatcaagaacaagggggtcattcaac 386
 Db 154188 GGGATGAGGAGGAAGAATGATGATGAAGTTATCAAGAACAAATGGGGGTCAATTCAC 154129
 QY 387 agttagatatctcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 431
 Db 154128 AGTTAGATATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 154084
 RESULT 13
 AC011450/c DNA PRI 29-APR-2000
 LOCUS Homo sapiens chromosome 19 clone CTC-30107, complete sequence.
 DEFINITION AC011450
 AC011450.4 GI:7670120
 VERSION
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 106168)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 106168)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (07-OCT-1999) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 3 (bases 1 to 106168)
 DOE Joint Genome Institute and Stanford Human Genome Center.
 Direct Submission
 Submitted (29-APR-2000) DOE Joint Genome Institute, 2800 Mitchell
 Drive, Walnut Creek, CA 94598, USA
 On Apr 29, 2000 this sequence version replaced gi:7021594.
 Draft Sequence Produced by DOE Joint Genome Institute
 www.jgi.doe.gov
 Finishing Completed at Stanford Human Genome Center
 www.sngc.stanford.edu
 Quality: Phrap Quality >=40 99.8% of Sequence;
 Estimated Total Number of Errors is 0.3.


```

/misc_feature      /note="assembly_name:Contig41
                   clone_end:SP6"
/misc_feature      /note="assembly_name:Contig42"
                   /note="assembly_name:Contig43"
misc_feature      /note="assembly_name:Contig43"
BASE COUNT      50350 a 40923 c 42172 g 43887 t 1401 others
ORIGIN
Query Match
Best Local Similarity 6.7%; Score 98; DB 73; Length 178733;
Matches 98; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 275 ggttccttggagggcgaagattcgtatcatatctccaaagaactgggaatg 334
    |||||||TCTTTAGGGCTCAGAGATTGCTGATATCATCTCCAAAGAACGGGAATGG 120161
Db 120162 GCTTCTCTTTAGGGCTCAGAGATTGCTGATATCATCTCCAAAGAACGGGAATGG 120161
QY 335 aggaagaagatgtgattgaattatcaggaacaaacg 372
    |||||||TCTTTAGGGCTCAGAGATTGCTGATATCATCTCCAAAGAACGGGAATGG 120161
Db 120162 AGGAAGAAGATGTGATTGAATTATCAGGAACAAACG 120199

```

```

RESULT 15
AC064836/c
LOCUS      AC064836      181861 bp      DNA      HTG      17-AUG-2000
DEFINITION Homo sapiens chromosome 2 clone RP11-68606, WORKING DRAFT SEQUENCE,
            17 unordered pieces.
ACCESSION  AC064836
VERSION     AC064836.4 GI:9838313
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
            1 (bases 1 to 181861)
            Waterston, R.H.
            The sequence of Homo sapiens clone
            Unpublished
            2 (bases 1 to 181861)
            Waterston, R.H.
            Direct Submission
            Submitted (22-APR-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
            On Aug 17, 2000 this sequence version replaced gi:7801500.

```

```

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Project Information
Center project name: H.NH0686006
Summary Statistics
Sequencing vector: M13, 100%
Sequencing vector: plasmid, 0%
Chemistry: Dye-primer ET, 100% of reads
Chemistry: Dye-terminator Big Dye, 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 173047 bases at least Q40
Consensus quality: 176179 bases at least Q30
Consensus quality: 178311 bases at least Q20
Insert size: 191000; agarose-fp
Quality coverage: 5.14 in Q20 bases; agarose-fp
Quality coverage: 5.45 in Q20 bases; sum-of-coverage
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence

```

```

FEATURES
source
* as soon as it is available and the accession number will
* be preserved.
1       1669: contig of 1669 bp in length
1670    1770: gap of unknown length
1770    3011: contig of 1242 bp in length
3011    3112: gap of unknown length
3112    5912: contig of 2801 bp in length
5912    6013: gap of unknown length
6013    7803: contig of 1790 bp in length
7803    7903: gap of unknown length
7903    12402: contig of 4499 bp in length
12402   12502: gap of unknown length
12502   16035: gap of 3533 bp in length
16035   16135: gap of unknown length
16135   21575: contig of 5440 bp in length
21575   21675: gap of unknown length
21675   26187: contig of 4512 bp in length
26187   26287: gap of unknown length
26287   31933: contig of 5646 bp in length
31933   32033: gap of unknown length
32033   40013: contig of 7980 bp in length
40013   40113: gap of unknown length
40113   46866: contig of 6753 bp in length
46866   46966: gap of unknown length
46966   56137: contig of 9172 bp in length
56137   56238: gap of unknown length
56238   64075: contig of 7738 bp in length
64075   64076: gap of unknown length
64076   78167: contig of 14092 bp in length
78167   78267: gap of unknown length
78267   95028: contig of 16761 bp in length
95028   95128: gap of unknown length
95128   127606: contig of 32478 bp in length
127606  127707: gap of unknown length
127707  181861: contig of 54155 bp in length.
Location/Qualifiers
1. 181861
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="2"
/clone="RP11-68606"
1. 1669
/note="assembly_name:Contig16"
1770. 3011
/note="assembly_name:Contig17"
3112. 5912
/note="assembly_name:Contig19"
6013. 7802
/note="assembly_name:Contig20"
7903. 12401
/note="assembly_name:Contig21"
12502. 16034
/note="assembly_name:Contig22"
16135. 21574
/note="assembly_name:Contig23"
21675. 26186
/note="assembly_name:Contig24"
26287. 31932
/note="assembly_name:Contig25"
32033. 40012
/note="assembly_name:Contig26"
40113. 46865
/note="assembly_name:Contig27"
46966. 56137
/note="assembly_name:Contig28"
56238. 63975
/note="assembly_name:Contig29
clone_end:T7
vector_side:right"
64076. 78167
/note="assembly_name:Contig30
clone_end:SP6
vector_side:left"

```

```

misc_feature      /note="assembly_name:Contig16"
misc_feature      /note="assembly_name:Contig17"
misc_feature      /note="assembly_name:Contig19"
misc_feature      /note="assembly_name:Contig20"
misc_feature      /note="assembly_name:Contig21"
misc_feature      /note="assembly_name:Contig22"
misc_feature      /note="assembly_name:Contig23"
misc_feature      /note="assembly_name:Contig24"
misc_feature      /note="assembly_name:Contig25"
misc_feature      /note="assembly_name:Contig26"
misc_feature      /note="assembly_name:Contig27"
misc_feature      /note="assembly_name:Contig28"
misc_feature      /note="assembly_name:Contig29
clone_end:T7
vector_side:right"
misc_feature      /note="assembly_name:Contig30
clone_end:SP6
vector_side:left"

```

Mon Aug 20 11:48:00 2001

us-09-484-964-1.oli.rge

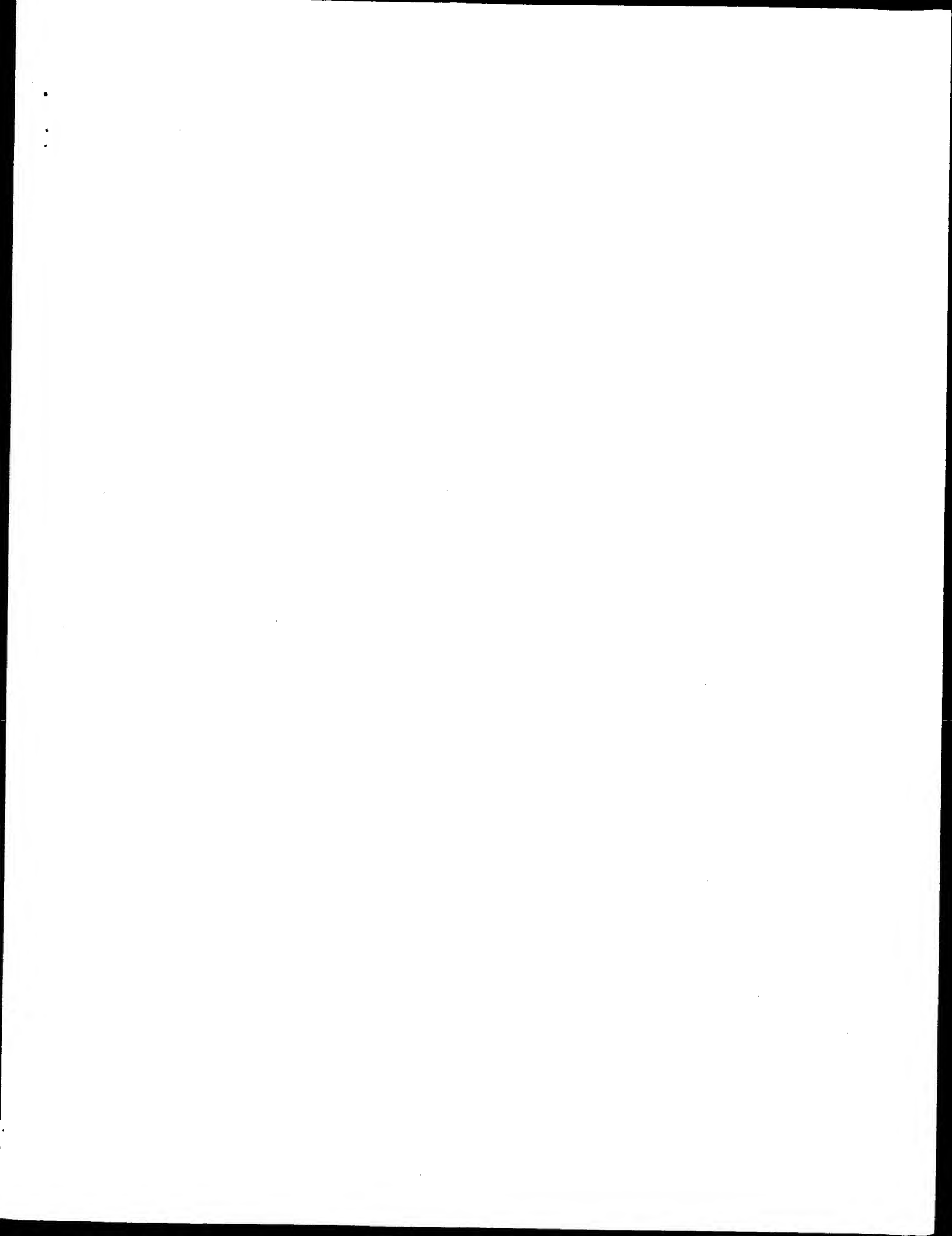
Page 13

misc_feature 78268 . 95028 /note="assembly_name:Contig31"
misc_feature 95129 . 127606 /note="assembly_name:Contig32"
misc_feature 127707 . 181861 /note="assembly_name:Contig33"
BASE COUNT 52240 a 39002 c 38693 g 50312 t 1614 others
ORIGIN

Query Match 6.6%; Score 97; DB 72; Length 181861;
Best Local Similarity 100.0%; Pred. No. 1.9e-37;
Matches 97; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 gaagatttgaaccgagcgaggttctgtaccgagcgctgtgtgcgaga 63
|||||
DB 71306 GAAGATTGTGTAACCCGAGCGAGGTTCTGTACCCGAGGCCCTGCTGTGCGAGA 71247
|||||
QY 64 ccccggtgaagccacgcacatcatgtgtgacag 100
|||||
DB 71246 CCCC GGGTGAAGCCACCGTCATCATGTGTGACGAG 71210
|||||

Search completed: August 17, 2001, 11:25:44
Job time: 6844 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2001, 09:34:05 ; Search time 84.61 Seconds
(without alignments)

3277.873 Million cell updates/sec

Title: US-09-484-964-1

Perfect score: 1465
Sequence: 1 cgggaagattgttaacc.....gataatggcgcgcagcc 1465

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 324599 seqs, 94655562 residues

Word size : 50

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	620	42.3	1514	2	US-09-213-768-1

ALIGNMENTS

QY	2	gggaagattgttaaccggaggttcgttaaccggagccgctgtgtcga 61
DB	50	gggaagattgttaaccggaggttcgttaaccggagccgctgtgtcga 109
QY	62	gaccccggtggaagccacgtcatctgtgtgacgaggaacacgtcaatgag 121
DB	110	gaccccggtggaagccacgtcatctgtgtgacgaggaacacgtcaatgag 169
QY	122	acttggggataagaagcaggtgataataataacccaagtcattggcaggatga 181
DB	170	acttggggataagaagcaggtgataataataacccaagtcattggcaggatga 229
QY	182	gtgagattcattcaagtgaaatgacaacacatctcaagaactcaagaatcact 241
DB	230	gtgagattcattcaagtgaaatgacaacacatctcaagaactcaagaatcact 289
QY	242	gtcaagaacaggtgttccaaatgaattcaactcaagttctcttgagggtgaagaattg 301
DB	290	gtcaagaacaggtgttccaaatgaattcaactcaagttctcttgagggtgaagaattg 349
QY	302	ctgataatcaatcccaaaagacgtgggaatggaggaagatgtgtgaattatc 361
DB	350	ctgataatcaatcccaaaagacgtgggaatggaggaagatgtgtgaattatc 409
QY	362	aggaacaaacgggggtcattcaacagtttagatatcttcttcttcttcc 421
DB	410	aggaacaaacgggggtcattcaacagtttagatatcttcttcttcttcc 469
QY	422	taatcttttttttttaaaatagttcttctgttaatgtgtgttcaaaacggaattg 481
DB	470	taatcttttttttttttaaaatagttcttctgttaatgtgtgttcaaaacggaattg 529
QY	482	aaaactgaccccatctctttgaaacatctgttaattgaaatctagtcattatc 541
DB	530	aaaactgaccccatctctttgaaacatctgttaattgaaatctagtcattatc 589
QY	542	attatgtttgttttcaatgtgtctgattttgtgtatcaagccatgacctcaatt 601
DB	590	attatgtttgttttcaatgtgtctgattttgtgtatcaagccatgacctcaatt 649
QY	602	acccttcctttttaaataatgagtgacagagagagtcacattttcaggaattga 661
DB	650	acccttcctttttaaataatgagtgacagagagagtcacattttcaggaattga 709
QY	662	ttttcaggtt 672
DB	710	ttttcaggtt 720

Search completed: August 17, 2001, 11:20:01
Job time: 6356 sec

RESULT 1
US-09-213-768-1
; Sequence 1, Application US/09213768
; Patent No. 5985664
; GENERAL INFORMATION:
; APPLICANT: Brenda F. Baker
; APPLICANT: Lex M. Cowsett
; TITLE OF INVENTION: ANTISENSE MODULATION OF SENTRIN EXPRESSION
; FILE REFERENCE: RTS-0026
; CURRENT APPLICATION NUMBER: US/09/213,768
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 47
; SEQ ID NO 1
; LENGTH: 1514
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (136)..(441)
US-09-213-768-1

Mon Aug 20 11:47:46 2001

us-09-484-964-1.oli.rni

Page 2

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2001, 10:09:15 ; Search time 131.18 Seconds
(without alignments)
7012.318 Million cell updates/sec

Title: US-09-484-964-1

Perfect score: 1465
Sequence: 1 cgggaagattgttaaaccc.....gataatggcatgcacgc 1465

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 730101 seqs, 313950809 residues

Word size : 50

Total number of hits satisfying chosen parameters: 9

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

N.Geneseq.0601.*
1: /SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1465	100.0	1465	19	AAV34564
2	895	61.1	1196	19	AAV59722
3	895	61.1	1220	19	AAV59557
4	620	42.3	1514	21	AA35859
5	300	20.5	616	21	AA16215
6	284	19.4	372	18	AA63339
7	284	19.4	372	22	AA59399
8	197	13.4	425	17	AA703735
9	85	5.8	180	16	AA119668

ALIGNMENTS

RESULT	1
ID	AAV34564
AC	AAV34564 standard; DNA; 1465 BP.
XX	
AC	AAV34564:
XX	
DT	11-SEP-1998 (first entry)
XX	
DE	Homo sapiens sentrin-1 gene.
XX	
KW	sentrin-1; protection; tumour necrosis factor; tnfi; apoptosis;
KW	Fas/Apo-Induced; tumour cell death; induction; tumour aggressiveness;
KW	detection; determination; ss.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	Location/Qualifiers
FT	CDS
FT	88..393
FT	/*tag= a
FT	/product= sentrin-1 polypeptide
XX	
PN	MO9820038-AL.
XX	
PD	14-MAY-1998.
XX	
PF	05-NOV-1997; 97WO-US20344.
XX	
PR	05-NOV-1996; 96US-0030302.
XX	
PA	(TEXA) UNIV TEXAS SYSTEM.
XX	
PI	Yeh ETH;
XX	
DR	WPI: 1998-286868/25.
DR	P-PSDB; AAM60079.
PT	New isolated sentrin polypeptide(s) - which inhibit TNF receptor or
PT	Fas/Apo-induced apoptosis, used to develop products for inducing
PT	cell death in tumours
XX	
PS	Claim 10; Page 76-77; 120pp; English.
XX	
CC	The sequence is that encoding the sentrin-1 polypeptide.
CC	Sentin polypeptides have the ability to protect or guard
CC	cells from tumour necrosis factor (TNF) or Fas/Apo induced
CC	cell death (apoptosis). Inhibitors of the sentrin polypeptides,
CC	e.g. antibodies, can be used for inducing cell death.
CC	particularly in tumours. The products can also be used for
CC	determining the aggressiveness of a tumour and for detection and
CC	isolation of products. The sentrin polypeptide can also be used to
CC	detect a ubiquitin conjugating enzyme polypeptide or PML polypeptide.
XX	
SQ	Sequence 1465 BP; 455 A; 248 C; 292 G; 470 T; 0 other;
XX	
Query Match	100.0%; Score 1465; DB 19; Length 1465;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 1465; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 cgggaagattgttaaaccccgagagcgttctgttaccgagcgcgtctgtgag 60
DB	1 cgggaagattgttaaaccccgagagcgttctgttaccgagcgcgtctgtgag 60
QY	61 agaccgccgggtgaagccacgcgtcatctgtcagcagaggaagcaacttaactgag 120
DB	61 agaccgccgggtgaagccacgcgtcatctgtcagcagaggaagcaacttaactgag 120
QY	121 gacttgggggataagagcaagtgatatacttaactcaagtcattgagacaggtaac 180
DB	121 gacttgggggataagagcaagtgatatacttaactcaagtcattgagacaggtaac 180

QY 181 agtgaatctcctcaagtgaaatgacaacacatctcaagaaatcaagaatcac 240
| | | | |
Db 181 agtgaatctcctcaagtgaaatgacaacacatctcaagaaatcaagaatcac 240
QY 241 tgcataagacagggtgttccatgaatctcaggttctcttgagggtcagagatt 300
| | | | |
Db 241 tgcataagacagggtgttccatgaatctcaggttctcttgagggtcagagatt 300
QY 301 gctataatcaatcctcaaaagaacatgagaaatgagaaatgagaaatgagaaat 360
| | | | |
Db 301 gctataatcaatcctcaaaagaacatgagaaatgagaaatgagaaatgagaaat 360
QY 361 caagaaacacagggtgttccatgaatctcaggttctcttgagggtcagagatt 420
| | | | |
Db 361 caagaaacacagggtgttccatgaatctcaggttctcttgagggtcagagatt 420
QY 421 ctaatcctcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 480
| | | | |
Db 421 ctaatcctcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 480
QY 481 gaaacacgaccccatctcttgaacacatctgtaattgtaattgtaattgtaatt 540
| | | | |
Db 481 gaaacacgaccccatctcttgaacacatctgtaattgtaattgtaattgtaatt 540
QY 541 cattattgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 600
| | | | |
Db 541 cattattgttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 600
QY 601 taccctctcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 660
| | | | |
Db 601 taccctctcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 660
QY 661 atttcaagctgt 720
| | | | |
Db 661 atttcaagctgt 720
QY 721 aattggcctgt 780
| | | | |
Db 721 aattggcctgt 780
QY 781 gatgtgaagtttctcagagacatgacatgacatgacatgacatgacatgacatgac 840
| | | | |
Db 781 gatgtgaagtttctcagagacatgacatgacatgacatgacatgacatgacatgac 840
QY 841 ctttaaaatttgagagctgtgacacacacacacacacacacacacacacacacacac 900
| | | | |
Db 841 ctttaaaatttgagagctgtgacacacacacacacacacacacacacacacacacac 900
QY 901 cagataaggtgagagatgactaactcacaagagagagagagagagagagagagagag 960
| | | | |
Db 901 cagataaggtgagagatgactaactcacaagagagagagagagagagagagagagag 960
QY 961 aagatttttcaaaaatctgtcagagagatcccaagaaatgcttaatttcaatgacat 1020
| | | | |
Db 961 aagatttttcaaaaatctgtcagagagatcccaagaaatgcttaatttcaatgacat 1020
QY 1021 aataaagctataactgacgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1080
| | | | |
Db 1021 aataaagctataactgacgaatgaatgaatgaatgaatgaatgaatgaatgaatgaat 1080
QY 1081 gtacttttggcccgagagatgagttttaaagtgacatctcttaccacatcttaaa 1140
| | | | |
Db 1081 gtacttttggcccgagagatgagttttaaagtgacatctcttaccacatcttaaa 1140
QY 1141 ataaacataatttgcacaaatcgtactaatttatttatttatttatttatttatttatt 1200
| | | | |
Db 1141 ataaacataatttgcacaaatcgtactaatttatttatttatttatttatttatttatt 1200
QY 1201 aaaaaatgctaaataaggttctctgtgacataatctgaaatctgacatggtacaat 1260
| | | | |
Db 1201 aaaaaatgctaaataaggttctctgtgacataatctgaaatctgacatggtacaat 1260

QY 1261 ttttctcattactgtacagagatgtttaaattgaatttggagcactgaaatgactga 1320
| | | | |
Db 1261 ttttctcattactgtacagagatgtttaaattgaatttggagcactgaaatgactga 1320
QY 1321 agtgccttctgaatcaaggaatttaagggcacaatacttctttaaactgactg 1380
| | | | |
Db 1321 agtgccttctgaatcaaggaatttaagggcacaatacttctttaaactgactg 1380
QY 1381 tttttttttaaacttgatatactccgtatgtgtcatatttgaacaggtacccaatca 1440
| | | | |
Db 1381 tttttttttaaacttgatatactccgtatgtgtcatatttgaacaggtacccaatca 1440
QY 1441 tgttgataaatgggcatgacagcc 1465
| | | | |
Db 1441 tgttgataaatgggcatgacagcc 1465
RESULT 2
AAV59722 standard; DNA; 1196 BP.
ID AAV59722 standard; DNA; 1196 BP.
XX
AC AAV59722;
XX
DT 19-JAN-1999 (first entry)
XX
DE Human secreted protein gene 47 clone HOGAV75.
XX
KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukemia;
developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
inflammation; ischemic shock; Alzheimer's disease; restenosis; AIDS;
cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX
OS Homo sapiens.
XX
PN WC9839448-A2.
XX
PD 11-SEP-1998.
XX
PF 06-MAR-1998; 98MO-US04493.
XX
PR 02-OCT-1997; 97US-006160.
PR 07-MAR-1997; 97US-0038621.
PR 07-MAR-1997; 97US-0040161.
PR 07-MAR-1997; 97US-0040162.
PR 07-MAR-1997; 97US-0040163.
PR 07-MAR-1997; 97US-0040333.
PR 07-MAR-1997; 97US-0040334.
PR 07-MAR-1997; 97US-0040336.
PR 07-MAR-1997; 97US-0040626.
PR 11-APR-1997; 97US-0043311.
PR 11-APR-1997; 97US-0043312.
PR 11-APR-1997; 97US-0043313.
PR 11-APR-1997; 97US-0043314.
PR 11-APR-1997; 97US-0043568.
PR 11-APR-1997; 97US-0043569.
PR 11-APR-1997; 97US-0043576.
PR 11-APR-1997; 97US-0043578.
PR 11-APR-1997; 97US-0043580.
PR 11-APR-1997; 97US-0043669.
PR 11-APR-1997; 97US-0043670.
PR 11-APR-1997; 97US-0043671.
PR 11-APR-1997; 97US-0043672.
PR 11-APR-1997; 97US-0043674.
PR 23-MAY-1997; 97US-0047492.
PR 23-MAY-1997; 97US-0047500.
PR 23-MAY-1997; 97US-0047501.
PR 23-MAY-1997; 97US-0047502.
PR 23-MAY-1997; 97US-0047503.
PR 23-MAY-1997; 97US-0047504.
PR 23-MAY-1997; 97US-0047581.

23-MAY-1997; 97US-0047582.
 23-MAY-1997; 97US-0047583.
 23-MAY-1997; 97US-0047584.
 23-MAY-1997; 97US-0047585.
 23-MAY-1997; 97US-0047586.
 23-MAY-1997; 97US-0047587.
 23-MAY-1997; 97US-0047588.
 23-MAY-1997; 97US-0047589.
 23-MAY-1997; 97US-0047590.
 23-MAY-1997; 97US-0047592.
 23-MAY-1997; 97US-0047593.
 23-MAY-1997; 97US-0047594.
 23-MAY-1997; 97US-0047595.
 23-MAY-1997; 97US-0047596.
 23-MAY-1997; 97US-0047597.
 23-MAY-1997; 97US-0047598.
 23-MAY-1997; 97US-0047599.
 23-MAY-1997; 97US-0047600.
 23-MAY-1997; 97US-0047601.
 23-MAY-1997; 97US-0047612.
 23-MAY-1997; 97US-0047613.
 23-MAY-1997; 97US-0047614.
 23-MAY-1997; 97US-0047615.
 23-MAY-1997; 97US-0047617.
 23-MAY-1997; 97US-0047618.
 23-MAY-1997; 97US-0047632.
 23-MAY-1997; 97US-0047633.
 06-JUN-1997; 97US-0048964.
 06-JUN-1997; 97US-0048974.
 13-JUN-1997; 97US-0049610.
 08-JUL-1997; 97US-0051926.
 18-JUL-1997; 97US-0052874.
 22-AUG-1997; 97US-0055724.
 22-AUG-1997; 97US-0056630.
 22-AUG-1997; 97US-0056631.
 22-AUG-1997; 97US-0056632.
 22-AUG-1997; 97US-0056636.
 22-AUG-1997; 97US-0056637.
 22-AUG-1997; 97US-0056662.
 22-AUG-1997; 97US-0056664.
 22-AUG-1997; 97US-0056845.
 22-AUG-1997; 97US-0056862.
 22-AUG-1997; 97US-0056864.
 22-AUG-1997; 97US-0056872.
 22-AUG-1997; 97US-0056874.
 22-AUG-1997; 97US-0056875.
 22-AUG-1997; 97US-0056876.
 22-AUG-1997; 97US-0056877.
 22-AUG-1997; 97US-0056878.
 22-AUG-1997; 97US-0056879.
 22-AUG-1997; 97US-0056880.
 22-AUG-1997; 97US-0056881.
 22-AUG-1997; 97US-0056882.
 22-AUG-1997; 97US-0056884.
 22-AUG-1997; 97US-0056886.
 22-AUG-1997; 97US-0056887.
 22-AUG-1997; 97US-0056888.
 22-AUG-1997; 97US-0056889.
 22-AUG-1997; 97US-0056892.
 22-AUG-1997; 97US-0056893.
 22-AUG-1997; 97US-0056894.
 22-AUG-1997; 97US-0056903.
 22-AUG-1997; 97US-0056908.
 22-AUG-1997; 97US-0056909.
 22-AUG-1997; 97US-0056910.
 22-AUG-1997; 97US-0056911.
 05-SEP-1997; 97US-0057650.
 05-SEP-1997; 97US-0057659.
 05-SEP-1997; 97US-0057761.
 12-SEP-1997; 97US-0058785.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX

PI Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GN;
 PI Feng P, Ferrie AM, Fischer CL, Florence KA, Greene JM, Hu JS;
 PI Kyaw H, Lallieur DM, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
 PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;
 XX WPI: 1998-506364/43.
 DR P-PSDB; AAW74938.
 XX
 XX New isolated human genes and the secreted polypeptide(s) they encode
 PT - useful for diagnosis and treatment of e.g. cancers, neurological
 PT disorders, immune diseases, inflammation or blood disorders
 XX
 PS Claim 1; Page 449; 721pp; English.
 XX
 XX This sequence represents a nucleic acid molecule designated Gene 47 from
 CC the human cDNA clone HOGAV75 (deposited as clone ATCC 97897 and ATCC
 CC 209043) which encodes a secreted human protein. The gene can be used to
 CC generate fusion proteins by linking to the gene to a human
 CC immunoglobulin Fc portion (e.g. AAV59502) for increasing the stability of
 CC the fused protein as compared to the human protein only.
 CC The invention relates to 186 novel genes and their fragments (nucleic
 CC acid sequences: AAV59511-V59612; amino acid sequences AAW74731-W75026)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological
 CC conditions can be diagnosed by determining the amount of the new
 CC polypeptides in a sample or by determining the presence of mutations in
 CC the new polynucleotides. Specific uses are described for each of the 186
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAV59511 for described uses).
 XX
 XX Sequence 1196 BP; 360 A; 212 C; 259 G; 364 T; 1 other;
 SO
 Query Match 61.1%; Score 895; DB 19; Length 1196;
 Best Local Similarity 99.8%; Pred. No. 0; Mismatches 1; Indels 1; Gaps 1;
 Matches 1065; Conservative 0;
 QY 4 gaaggaattgtaaaccccgagaggtctgctaccgagcgccgtctgctgagaga 63
 |||||||
 DB 50 gaaggaattgtaaaccccgagaggtctgctaccgagcgccgtctgctgagaga 109
 |||||||
 QY 64 ccccgaggtgaagccacgctcatatgtctgcacgagggcaaaaccttaactgaagac 123
 |||||||
 DB 110 ccccgaggtgaagccacgctcatatgtctgcacgagggcaaaaccttaactgaagac 169
 |||||||
 QY 124 ttgggggataaagaagatgaatataataataactcaagtcattgagcaagatagcagt 183
 |||||||
 DB 170 ttgggggataaagaagatgaatataataataactcaagtcattgagcaagatagcagt 229
 |||||||
 QY 184 gagattcaactcaaatgaaatgaaacacacatctcaagaagaactcaagaatcatatcgt 243
 |||||||
 DB 230 gagattcaactcaaatgaaatgaaacacacatctcaagaagaactcaagaatcatatcgt 289
 |||||||
 QY 244 caaagaacgggtgtcccaatgaattcaaccacggttctcttgaagggccaagaattgct 303
 |||||||
 DB 290 caaagaacgggtgtcccaatgaattcaaccacggttctcttgaagggccaagaattgct 349
 |||||||
 QY 304 gataatcatatctcaaaagaacgtgggaatggaagaagaatgattgaattatcaag 363
 |||||||
 DB 350 gataatcatatctcaaaagaacgtgggaatggaagaagaatgattgaattatcaag 409
 |||||||
 QY 364 gaacaaacgggggtgctatcaacagtttagatactctttatctttcttcttcctc 423
 |||||||
 DB 410 gaacaaacgggggtgctatcaacagtttagatactctttatctttcttcttcctc 469
 |||||||
 QY 424 aatccttttttttttaaaatagttcttttgaatggtgtgttcaaaacgaattgaa 483
 |||||||
 DB 470 aatccttttttttttaaaatagttctttgaaatggtgtgttcaaaacgaattgaa 529
 |||||||
 QY 484 aactggaaccccatctcttgaacacatctgttaattgaaattcagtcattatcat 543
 |||||||
 DB 530 aactggaaccccatctcttgaacacatctgttaattgaaattcagtcattatcat 589
 |||||||

[illegible]

PR 23-MAY-1997; 97US-0047600
PR 23-MAY-1997; 97US-0047601
PR 23-MAY-1997; 97US-0047612

23-MAY-1997; 97US-0047613

PR	23-MAY-1997;	97US-0047615
PR	23-MAY-1997;	97US-0047617

PR	23-MAY-1997:	97US-0047632
PR	23-MAY-1997:	97US-0047618

PRR	06-JUN-1997	9705-00489664
PRR	06-JUN-1997	9705-0048974
PRR	13-JUN-1997	9705-0049610
PRR	06-JUL-1997	9705-0051926
PRR	18-JUL-1997	9705-0052874
PRR	18-AUG-1997	9705-0055724
PRR	22-AUG-1997	9705-0056630
PRR	22-AUG-1997	9705-0056631

22-AUG-1997; 97US-0056632

22-AUG-1997; 97US-0056637

22-AUG-1997; 97US-0056664

22-AUG-1997; 97US-0056862

22-AUG-1997; 97US-0056864

22-AUG-1997; 97US-0056874.

22-AUG-1997; 97US-0056876.

22-AUG-1997; 97US-0056878

22-AUG-1997; 9705-0056879.
22-AUG-1997; 9705-0056880

22-AUG-1997; 97US-0056881.

1000

22-AUG-1997; 97US-0056884.
 PR 22-AUG-1997; 97US-0056886.
 PR 22-AUG-1997; 97US-0056887.
 PR 22-AUG-1997; 97US-0056888.
 PR 22-AUG-1997; 97US-0056889.
 PR 22-AUG-1997; 97US-0056892.
 PR 22-AUG-1997; 97US-0056893.
 PR 22-AUG-1997; 97US-0056894.
 PR 22-AUG-1997; 97US-0056903.
 PR 22-AUG-1997; 97US-0056908.
 PR 22-AUG-1997; 97US-0056909.
 PR 22-AUG-1997; 97US-0056910.
 PR 22-AUG-1997; 97US-0056911.
 PR 05-SEP-1997; 97US-0057650.
 PR 05-SEP-1997; 97US-0057659.
 PR 05-SEP-1997; 97US-0057761.
 PR 12-SEP-1997; 97US-0058785.

(HUMA-) HUMAN GENOME SCI INC.

XX Bednarik DP, Brewer LA, Carter KC, Duan R, Ebner R, Endress GA;
 PI Feng P, Ferlie AM, Fischer CL, Florence KA, Greene JM, Hu JS;
 PI Kyaw H, Lafleur DW, Li Y, Moore PA, Ni J, Olsen HS, Rosen CA;
 PI Ruben SM, Shi Y, Soppet DR, Young PE, Yu GL, Zeng Z;

WPI: 1998-506364/43.
 P-PSDB; AAM74777.

XX New isolated human genes and the secreted polypeptide(s) they encode
 PT - useful for diagnosis and treatment of e.g. cancers, neurological
 disorders, immune diseases, inflammation or blood disorders

PS Claim 1: Page 279-280; 721pp; English.

XX This sequence represents a nucleic acid molecule designated Gene 47 from
 CC the human CDNA clone HOGAV75 (deposited as clone ATCC 97899 and ATCC
 CC 209045) which encodes a secreted human protein. The gene can be used to
 CC generate fusion proteins by linking to the gene to a human
 CC immunoglobulin Fc portion (e.g. AAV59502) for increasing the stability of
 CC the fused protein as compared to the human protein only.
 CC The invention relates to 186 novel genes and their fragments (nucleic
 CC acid sequences: AAV59511-V59812; amino acid sequences AAM74731-W5026)
 CC which are useful for preventing, treating or ameliorating medical
 CC conditions e.g. by protein or gene therapy. Also, pathological conditions
 CC can be diagnosed by determining the amount of the new polypeptides in a
 CC sample or by determining the presence of mutations in the new
 CC polynucleotides. Specific uses are described for each of the 186
 CC polynucleotides, based on which tissues they are most highly expressed in
 CC (see AAV59511 for described uses).

XX Sequence 1220 BP; 382 A; 207 C; 258 G; 372 T; 1 other;

Query Match 61.1%; Score 895; DB 19; Length 1220;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1063; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 4 gaagatttgaaccccgagcgaggttctgtcttaaccggagccgctgtctgcgaga 63
 DB 44 gaagatttgaaccccgagcgaggttctgtcttaaccggagccgctgtctgcgaga 103
 QY 64 ccccggttgaagccacgctcatctgtctgacagagggagcaaaccttaactgaagac 123
 DB 104 ccccggttgaagccacgctcatctgtctgacagagggagcaaaccttaactgaagac 163
 QY 124 ttgggggataaagaacgaagtgaaataataactcaagtcattgtgacaggaatagcagt 183
 DB 164 ttgggggataaagaagtgaaataataactcaagtcattgtgacaggaatagcagt 223
 QY 184 gagattcaactcaaaagtgaatgaacacacatctccaagaagaactaaagaatctactgt 243
 DB 224 gagattcaactcaaaagtgaatgaacacacatctccaagaagaactaaagaatctactgt 283

QY 244 caaagacaggtgttccaatgaatctcaaggttctctttagagggtcagaagatgct 303
 DB 284 caaagacaggtgttccaatgaatctcaaggttctctttagagggtcagaagatgct 343
 QY 304 gataataactccaagaagacttggaatgaggaagaagatgtagatgtagatgtag 363
 DB 344 gataataactccaagaagacttggaatgaggaagaagatgtagatgtagatgtag 403
 QY 364 gaacaacgggggggtcattcaacagtttagatctcttatttttttttttttttttt 423
 DB 404 gaacaacgggggggtcattcaacagtttagatctcttatttttttttttttttttt 463
 QY 424 aatccttt 483
 DB 464 aatccttt 523
 QY 484 aactggcaccatctcttgaacatctggttaaattgaattcagtgatctatctaat 543
 DB 524 aactggcaccatctcttgaacatctggttaaattgaattcagtgatctatctaat 583
 QY 544 tatgtttgttttcaatgtctgatttttgggtgacagctcctcagctccctcattatc 603
 DB 584 tatgtttgttttcaatgtctgatttttgggtgacagctcctcagctccctcattatc 643
 QY 604 cctctccttttaaaattacgtgtgacagagaggtcacttttcagagcattgcat 663
 DB 644 cctctccttttaaaattacgtgtgacagagaggtcacttttcagagcattgcat 703
 QY 664 ttcaaggttgggtgataataagaatcgaatgcaagtgatcattcaatgacttccat 723
 DB 704 ttcaaggttgggtgataataagaatcgaatgcaagtgatcattcaatgacttccat 763
 QY 724 tggccttgatgttc-actatgtgattacttctccttgacttgacttgacttgagga 782
 DB 764 tggccttgatgttc-actatgtgattacttctccttgacttgacttgacttgagga 823
 QY 783 tggagattttcagaagaactgacgttggaataatgaccttcttaacttgaagctact 842
 DB 824 tggagattttcagaagaactgacgttggaataatgaccttcttaacttgaagctact 883
 QY 843 tttaaaatttggaggtctgacccaagaagaaggaatcaggttgaagtcagaatgaca 902
 DB 884 tttaaaatttggaggtctgacccaagaagaaggaatcaggttgaagtcagaatgaca 943
 QY 903 gataaagtgaagtaatgacttaactccaagaatgcttcaactgaaagaagacatttaa 962
 DB 944 gataaagtgaagtaatgacttaactccaagaatgcttcaactgaaagaagacatttaa 1003
 QY 963 gatttttaaaaatctgttcagaagaatcccaagaagaagttcatttcaatgcaattaa 1022
 DB 1004 gatttttaaaaatctgttcagaagaatcccaagaagaagttcatttcaatgcaattaa 1063
 QY 1023 taaagctatacatgcaagaatgatatcaacaagaacactgctctttt 1069
 DB 1064 taaagctatacatgcaagaatgatatcaacaagaacactgctctttt 1110

RESULT 4

AA235859 standard; cDNA; 1514 BP.

AA235859;

03-FEB-2000 (first entry)

Human sentrin nucleotide sequence.

Human; sentrin; antisense oligonucleotide; phosphorothioate;

inhibition; modulation; expression; diagnosis; ss.

Homo sapiens.

Location/Qualifiers

Key

Best Local Similarity 99.7%; Pred. No. 1.2e-131;
Matches 350; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4 gaagatttgaaccccgagagaggttctgtcttaccgagccgtctgtcgagaga 63
DB 43 gaagatttgaaccccgagagaggttctgtcttaccgagccgtctgtcgagaga 102
QY 64 ccccggttgaagcccgatcatatgtctgacagaggaacaaacttcaactgagac 123
DB 103 ccccggttgaagcccgatcatatgtctgacagaggaacaaacttcaactgagac 162
QY 124 ttgggggataagaaagtgatgataatataactcaaatgcatgtgacagagatg 183
DB 163 ttgggggataagaaagtgatgataatataactcaaatgcatgtgacagagatg 222
QY 184 gagattcattcaaaagttaaaatgacacacatctcaagaactcaagaatcattgt 243
DB 223 gagattcattcaaaagttaaaatgacacacatctcaagaactcaagaatcattgt 282
QY 244 caaagacaggggttccaatgattcactcaggttctcttcttgagggtcagagatgtct 303
DB 283 caaagacaggggttccaatgattcactcaggttctcttcttgagggtcagagatgtct 342
QY 304 gataatcattcacaagaactgggaatgaggaagaagatgtgattgaa 354
DB 343 gataatcattcacaagaactgggaatgaggaagaagatgtgattgaa 393

RESULT 6
AAT63339
ID AAT63339 standard; CDNA; 372 BP.

XX AAT63339;

DT 17-AUG-1997 (first entry)

DE Human host cell protein NPI-6 partial CDNA clone.

KW Host cell protein; NPI-6; nucleoprotein interactor 6;

KW Influenza virus; replication; antiviral; virucide; ss.

XX Homo sapiens.

XX MOJ9712967-A1.

XX 10-APR-1997.

PF 06-OCT-1995; 95MO-US13044.

PR 06-OCT-1995; 95MO-US13044.

PA (MOJN) MOUNT SINAI MEDICAL CENT.

PI Oneill R. Palese P;

DR WPI; 1997-226211/20.

PT New isolated DNA which encodes viral interacting proteins - used in
PT assays to isolate products for inhibiting viral protein binding
PT which is required for infection, replication, assembly or release

PS Disclosure; Fig 11; 98pp; English.

XX CDNA clones (AAT63335-39) comprise partial sequences for human
XX nucleoprotein interactor proteins NPI-2 to NPI-6, respectively.
XX These are host cell proteins which interact with influenza virus
XX nucleoprotein (NP) and which may be accessory proteins required for
XX influenza virus replication. NPI-6 was identified as a novel
XX protein sequence. NPI sequences were isolated by interactive trap
XX selection using LexA-NP as bait and yeast transformed with an HeLa
XX CDNA library. NPI-1 (see also AAT63334) and NPI-2 to NPI-6 CDNAs were
XX cloned and analysed. The clones can be used to produce human NPI
XX proteins for use in identifying cpds. that inhibit viral

CC replication.

XX Sequence 372 BP; 121 A; 72 C; 97 G; 82 T; 0 other;

XX Sequence 372 BP; 121 A; 72 C; 97 G; 82 T; 0 other;

Query Match 19.4%; Score 284; DB 18; Length 372;

Best Local Similarity 99.7%; Pred. No. 4.2e-124;

Matches 334; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 atttgaaccccgagagaggttctgtcttaccgagccgtctgtcgagagacccc 68
DB 1 atttgaaccccgagagaggttctgtcttaccgagccgtctgtcgagagacccc 60
QY 69 gggtaagccacccgtcatatgtctgacagaggaacaaacttcaactgagacttg 128
DB 61 gggtaagccacccgtcatatgtctgacagaggaacaaacttcaactgagacttg 120
QY 129 gataagaagcaagtgataatataactcaaatgcatgtgacagagatgagat 188
DB 121 gataagaagcaagtgataatataactcaaatgcatgtgacagagatgagat 180
QY 189 tcacttcaaaagtgaatgacacacacatctcaagaactcaagaatcacaatgtgaa 248
DB 181 tcacttcaaaagtgaatgacacacacatctcaagaactcaagaatcacaatgtgaa 240
QY 249 acaaggtgttccaatgattcactcaggttctcttcttgagggtcagagatgtgaa 308
DB 241 acaaggtgttccaatgattcactcaggttctcttcttgagggtcagagatgtgaa 300
QY 309 tcacttcaaaagtgaatgacacacacatctcaagaactcaagaatcacaatgtgaa 343
DB 301 tcacttcaaaagtgaatgacacacacatctcaagaactcaagaatcacaatgtgaa 335

RESULT 7

AAF59399
ID AAF59399 standard; DNA; 372 BP.

XX AAF59399;

DT 02-MAY-2001 (first entry)

DE Human host cell protein NPI-6 partial nucleotide sequence SMQ ID NO.11.

XX Identification; antiviral; viral protein; viral replication; NP;

KW viral infection; nucleoprotein; ds.

XX Homo sapiens.

XX MO200111335-A2.

XX 15-FEB-2001.

PF 11-AUG-2000; 2000MO-US22257.

PR 11-AUG-1999; 9905-0148263.

PA (MOJN) MOUNT SINAI SCHOOL MEDICINE.

PI O'Neill R. Harty R. Palese PM;

DR WPI; 2001-168816/17.

PT Identifying a substance that inhibits the interaction between a viral
PT protein and a host cell protein, useful for the discovery of new
PT antiviral compounds -

PS Disclosure; Fig 11; 147pp; English.

XX The present invention describes a method (M1) for identifying a
XX substance that inhibits the interaction of a viral protein (VP) with a
XX host cell protein (HP). The method comprises: (a) contacting HP with VP
XX in the presence of a test substance; and (b) detecting complex formation,

CC where the ability of the test substance to inhibit HP/VP interaction is
 CC indicated by a decrease in complex formation. The antiviral compounds
 CC that inhibit the interaction between a host protein (NSI-BP or NPI-1)
 CC and a viral protein (NSI) are useful for treating or inhibiting viral
 CC infection, preferably influenza and rhinovirus infection, in humans.
 CC Antiviral compounds include peptides and antibodies. In particular
 CC compositions comprising a polypeptide containing an amino acid sequence
 CC corresponding to the NP-MS domain of the influenza virus NP protein,
 CC which inhibits the specific interaction of the NPI-1 protein with the
 CC influenza virus NP protein are useful for treating or inhibiting
 CC influenza viral infection in humans. The present sequence represents a
 CC human host cell protein NPI-6 partial nucleotide sequence, which is used
 CC in the exemplification of the present invention.
 CC
 CC Sequence 372 BP; 121 A; 72 C; 97 G; 82 T; 0 other;

Query Match 19.4%; Score 284; DB 22; Length 372;
 Best Local Similarity 99.7%; Pred. No. 4.2e-124;
 Matches 334; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 9 attgtgaaccccgagaggttctgtaccggagccgctgtgtcgagagacc 68
 Db 1 attgtgaaccccgagaggttctgtaccggagccgctgtgtcgagagacc 60
 QY 69 ggggtgaagccacccatcatgtctgtaccagagagcaaaccttcaactgaagattgg 128
 Db 61 ggggtgaagccacccatcatgtctgtaccagagagcaaaccttcaactgaagattgg 120
 QY 129 ggttaagagcaagtgatataataaactcaatgcatgtgacagagatgagat 188
 Db 121 ggttaagagcaagtgatataataaactcaatgcatgtgacagagatgagat 180
 QY 189 tcaactcaagtgaataatgacacacatctcaagaactcaagaatcatatgtcagaag 248
 Db 181 tcaactcaagtgaataatgacacacatctcaagaactcaagaatcatatgtcagaag 240
 QY 249 acaagggtgttcaatgaattcactcaggtttctcttgaaggtcagagaattgtcgataa 308
 Db 241 acaagggtgttcaatgaattcactcaggtttctcttgaaggtcagagaattgtcgataa 300
 QY 309 tcaactccaaaagaactcgtggaatgagagaag 343
 Db 301 tcaactccaaaagaactcgtggaatgagagaag 335

RESULT 8

AAT03735
 ID AAT03735 standard; cDNA; 425 BP.

AC AAT03735;

DT 26-MAR-1996 (first entry)

DE TNF-R p55IC/Fas-IC-binding protein DD11 cDNA clone.

KW Tumour necrosis factor receptor; TNF-R; p55IC; Fas-IC;
 intracellular domain binding protein; ss.

KM Homo sapiens.

OS Homo sapiens.

PN W09531544-A1.

PD 23-NOV-1995.

PF 11-MAY-1995; 95MO-US05854.

PR 02-OCT-1994; 94IL-0111125.

PA 11-MAY-1994; 94IL-0109632.

XX (WEIN/) WEINMURZEL H.
 PA (YEDA) YEDA RES & DEV CO LTD.
 XX

PI Boldin M, Melt I, Varfolomeev E, Wallach D;
 XX WPI; 1996-010930/01.
 DR
 XX
 PT TNF-NGF receptor superfamily intracellular domain-binding proteins -
 PT useful for modulating receptor function, e.g. for treating tumours
 PT or HIV-infected cells
 XX
 PS Claim 18; Fig 12; 96pp; English.
 CC
 CC cDNA clone DD11 (AAT03735) was isolated by screening a human Hela
 CC library using the human tumour necrosis factor receptor (TNF-R)
 CC p55 death domain (p55D) as 'bait'. The full length of the DD11
 CC transcript is approx 1.2 kb. The protein encoded by cDNA DD11
 CC interacts strongly with p55D (amino acids 326-414) and also with
 CC human and mouse FAS intracellular domain. Such proteins may
 CC be used modulate TNF-R function.
 CC
 CC Sequence 425 BP; 135 A; 80 C; 115 G; 95 T; 0 other;

Query Match 13.4%; Score 197; DB 17; Length 425;
 Best Local Similarity 100.0%; Pred. No. 4.1e-83;
 Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 140 aaggtgaatatataaactcaagatcatgtgacagagatagacagtgatcactcaag 199
 Db 171 aaggtgaatatataaactcaagatcatgtgacagagatagacagtgatcactcaag 230
 QY 200 tgaatatgacacacatctcaagaactcaagaatcatatcttcaagaaggtgttc 259
 Db 231 tgaatatgacacacatctcaagaactcaagaatcatatcttcaagaaggtgttc 290
 QY 260 caatgaattcactcaggtttctcttgaaggtcagagaattgtgataatcactca 319
 Db 291 caatgaattcactcaggtttctcttgaaggtcagagaattgtgataatcactca 350
 QY 320 aagaactgggaatggag 336
 Db 351 aagaactgggaatggag 367

RESULT 9

AAT19668
 ID AAT19668 standard; cDNA to mRNA; 180 BP.

AC AAT19668;

DT 28-JUN-1996 (first entry)

DE Human gene signature HDMG500738.

KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
 human; cloning; mapping; non-biased library; diagnosis; detection;
 cell typing; abnormal cell function; ss.

KM Homo sapiens.

OS Homo sapiens.

PN W09514772-A1.

PD 01-JUN-1995.

PF 11-NOV-1994; 94MO-JP01916.

PR 12-NOV-1993; 93JP-0355504.

PA (MATS/) MATSUBARA K.
 PA (OKUBO/) OKUBO K.

PI Matsubara K, Okubo K;

DR WPI; 1995-206931/27.

PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

PS Claim 1; Page 445; 2245pp; Japanese.

A single-stranded DNA (or its complementary strand or the corresponding double-stranded DNA) which comprises one of the 7837 "GS" sequences given in A0119001-126837 and which is able to hybridise to part of human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature) sequences were obtained from 3'-directed cDNA libraries prepared from various human tissues; synthesis of cDNA was initiated from the 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-untranslated sequence is unique to a particular mRNA species, almost all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library is constructed so as to reflect accurately the relative abundance of different mRNAs in the particular tissue from which it was derived. The appearance frequency of a given GS in a cDNA library can be determined (esp. using primers and probes derived from the GS sequences) as a means of diagnosing abnormal cell function or for recognising different cell types.

SQ Sequence 180 BP; 65 A; 26 C; 26 G; 63 T; 0 other;

Query Match	5.8%;	Score 85;	DB 16;	Length 180;
Best Local Similarity	100.0%;	Pred. No.	2.4e-30;	
Matches 85;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

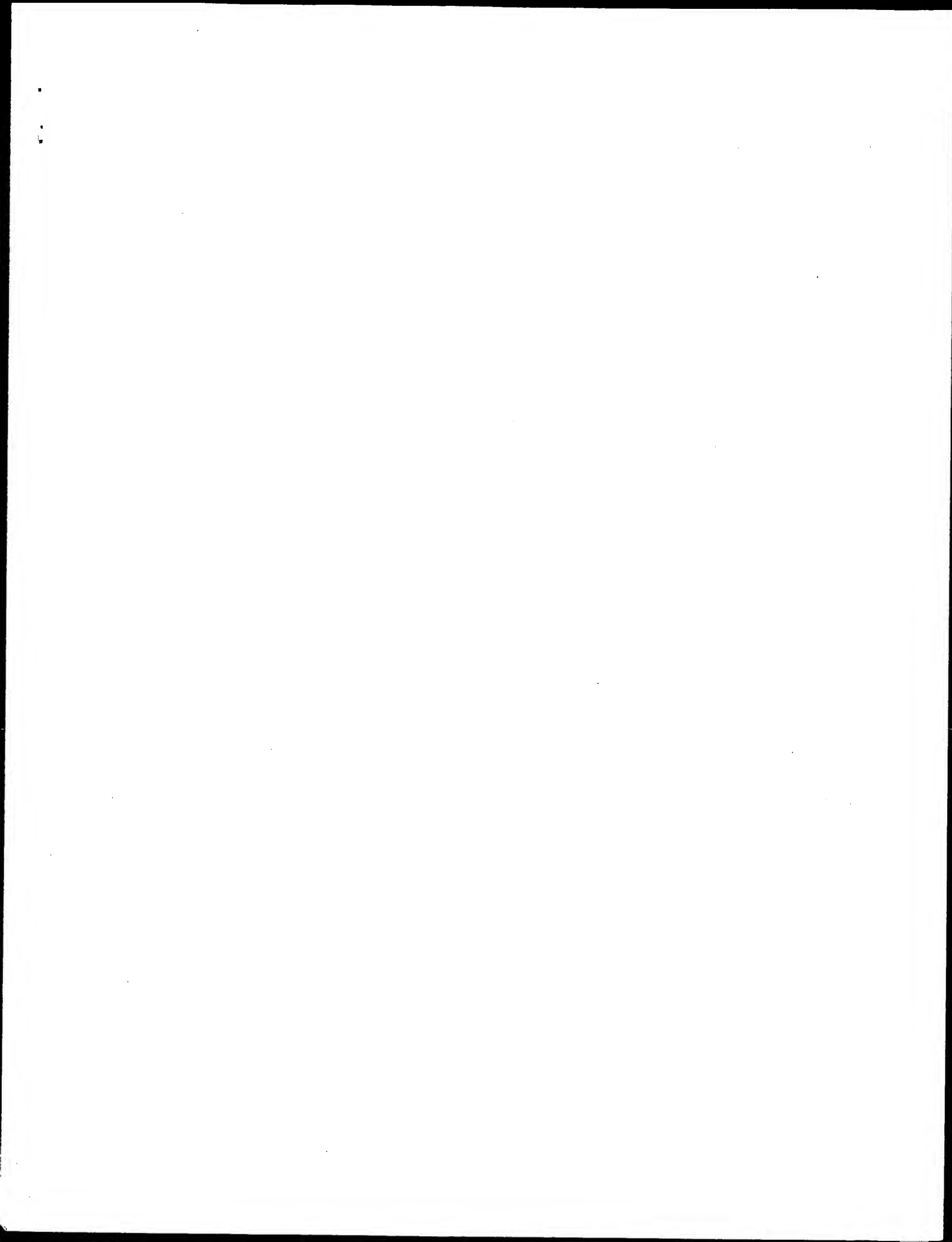
QY 1071 gattttattgtactttttgacctggatatgtggtttaatgacattgtctgtaccag 1130

Db 84 gatttattgtactttttgacctgataatggttttaatatgacattgtctgtaccag 143

QY 1131 cttcataaataacaatatgt 1155

Db 144 cttcattaataacaatatgt 168

Search completed: August 17, 2001, 11:22:29
Job time: 4394 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 17, 2001, 09:30:24 ; Search time 1171.13 Seconds
(without alignments)
11824.845 Million cell updates/sec

Title: us-09-484-964-1

Perfect score: 1465
Sequence: 1 cgggaagattgtaaaccc.....gataatggcattgccagcc 1465

Scoring table: OLIGO_NUC
Gapop 60.0 , Capext 60.0

Searched: 10228115 seqs, 4726426750 residues

Word size : 50

Total number of hits satisfying chosen parameters: 566

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

BST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_esthum29:*
64: em_esthum30:*
65: em_esthum31:*
66: em_esthum32:*
67: em_esthum33:*
68: em_esthum34:*
69: em_esthum35:*
70: em_esthum36:*
71: em_esthum37:*
72: em_esthum38:*
73: em_esthum39:*
74: em_esthum40:*
75: em_esthum41:*
76: em_esthum42:*
77: em_esthum43:*
78: em_esthum44:*
79: em_esthum45:*
80: em_esthum46:*
81: em_esthum47:*
82: em_esthum48:*
83: em_esthum49:*
84: em_esthum50:*
85: em_esthum51:*
86: em_esthum52:*
87: em_esthum53:*
88: em_esthum54:*
89: em_esthum55:*
90: em_esthum56:*
91: em_esthum57:*
92: em_esthum58:*
93: em_esthum59:*
94: em_esthum60:*
95: em_esthum61:*
96: em_esthum62:*
97: em_esthum63:*
98: em_esthum64:*
99: em_esthum65:*
100: em_esthum66:*
101: em_esthum67:*
102: em_esthum68:*
103: em_esthum69:*
104: em_esthum70:*
105: em_esthum71:*
106: em_esthum72:*
107: em_esthum73:*
108: em_esthum74:*
109: em_esthum75:*
110: em_esthum76:*
111: em_esthum77:*
112: em_esthum78:*
113: em_esthum79:*
114: em_esthum80:*
115: em_esthum81:*
116: em_esthum82:*

```
117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: em_est110:*
180: em_est111:*
181: em_est112:*
182: em_est113:*
183: em_est114:*
184: em_est115:*
185: em_est116:*
186: em_est117:*
187: em_est118:*
188: em_est119:*
189: em_est120:*
190: em_est121:*
191: em_est122:*
192: em_est123:*
193: em_est124:*
194: em_est125:*
195: em_est126:*
196: em_est127:*
197: em_est128:*
198: em_est129:*
199: em_est130:*
200: em_est131:*
201: em_est132:*
202: em_est133:*
203: em_est134:*
204: em_est135:*
205: em_est136:*
206: em_est137:*
207: em_est138:*
208: em_est139:*
209: em_est140:*
210: em_est141:*
211: em_est142:*
212: em_est143:*
213: em_est144:*
214: em_est145:*
215: em_est146:*
216: em_est147:*
217: em_est148:*
218: em_est149:*
219: em_est150:*
220: em_est151:*
221: em_est152:*
222: em_est153:*
223: em_est154:*
224: em_est155:*
225: em_est156:*
226: em_est157:*
227: em_est158:*
228: em_est159:*
229: em_est160:*
230: em_est161:*
231: em_est162:*
232: em_est163:*
233: em_est164:*
234: em_est165:*
235: em_est166:*
236: em_est167:*
237: em_est168:*
238: em_est169:*
239: em_est170:*
240: em_est171:*
241: em_est172:*
242: em_est173:*
243: em_est174:*
244: em_est175:*
245: em_est176:*
246: em_est177:*
247: em_est178:*
248: em_est179:*
249: em_est180:*
250: em_est181:*
251: em_est182:*
252: em_est183:*
253: em_est184:*
254: em_est185:*
255: em_est186:*
256: em_est187:*
257: em_est188:*
258: em_est189:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

/dev stage="pooled tissue from post conception fetuses (20 week, 24 week and 26 week)"
 /lab_host="DH10B"
 /note="Organ: Fetal brain; Vector: PCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville , Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 253 a 166 c 206 g 280 t 2 others

ORIGIN

Query Match 46.6%; Score 683; DB 106; Length 907;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 733; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

4 gaaggaattgtataaccccgagagaggtctgcttaaccgagccgctgctgctgagaga 63
 |||||||
 46 GAAGGATTTGTAAACCCCGAGCGAGGTTCTGCTTACCCGAGCGCTGCTGCGGAGAGA 105
 |||||||
 64 ccccggtggaagcaccgctcatctgctgacgagaggaacaccccttaactgagagac 123
 |||||||
 106 ccccggtggaagcaccgctcatctgctgacgagaggaacaccccttaactgagagac 165
 |||||||
 124 ttgggggataaagaagagtgataataataaactcaagtcattgagacagatagagcag 183
 |||||||
 166 TTGGGGGATTAAGACGAGAGTGATATATTAAACCTCAAGTCATTGAGACGATAGCAGT 225
 |||||||
 184 gagatctacttaaaagtgaataatgacacacatctcaagaacataaagaatcattgct 243
 |||||||
 226 GAGATTCACCTCAAAAGTAAATGACAACATCTCAAGAACTCAAGAAATCATATCTGT 285
 |||||||
 244 caaagaacgggtgtctcaatgactcactcaagttctccttgagagtcagagaattgct 303
 |||||||
 286 CAAGAACAGGGGTCTCCATGATGATCTCAAGTTCTCTTTGAGGGTCAGAGAAATTCCT 345
 |||||||
 304 gataatcactccaagaactgggaatggaagaagatgtagtgaattatcag 363
 |||||||
 346 GATATATCATACTCCAAAAGAACTCGGAATGAGCAAGATGATGATGATGATGATGAT 405
 |||||||
 364 gacacaaagggggggctcatcaacagttagatattcttttctttcttctcctc 423
 |||||||
 406 GAACAACAGGGGGGTCAATCAACAGTTAGATATCTTTTATTTTCTTTTCTTTCCCTC 465
 |||||||
 424 aatcctttttaaataaatagtctcttgaatgtagtgcacaaagaaatgaa 483
 |||||||
 466 AATCCTTTTATTTTAAAAAATAGTCTTTTGTATGTGGGTTCANAAACGAAATTTAA 525
 |||||||
 484 aactggcaccacatcttgaacatctgtaattgaattcagtgctatattatcat 543
 |||||||
 526 AACTGGCAACCCCATCTCTTTGAACATCTGGTAATTTGAATTTCTAGTCTCATATTCAT 585
 |||||||
 544 tatgtttgtttctatctgtgctgatttttggtgatacaagctcagtcctccatataac 603
 |||||||
 586 TATTTTGTGTTTCAATTTGCTGATTTTGGATTTTGGATCAAGCCCTCATCTCATATTTAC 645
 |||||||
 604 cctcccttttaaaataatcagtgtagcagagagagtcacccctttagggacattgact 663
 |||||||
 646 CCTTCCTTTTAAAAAATAGTCTGTCAGACAGAGAGTCACTTTTTCAGGACATTTGCATT 705
 |||||||
 664 ttcaaggtcttggtgataataaagtcgaaccaatgcaatgctcattcaatgaattccat 723
 |||||||
 706 TTCAGGCTTGTGATTAATTAAGATCGACCAATGCAAGTCTCATATGACTTTCCAAAT 765
 |||||||
 724 tggccctatgcttc 737
 |||||||
 766 tggccctatgcttc 779

RESULT 4
 AL583458 913 bp mRNA EST 16-FEB-2001
 LOCUS AL583458 LTI_NFL010_BC2 Homo sapiens cDNA clone CS0DL012YD05 5
 DEFINITION prime, mRNA sequence.
 ACCESSION AL583458
 VERSION AL583458.1 GI:12952440
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 REFERENCE 1 (bases 1 to 913)
 AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polyes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT Contact: Genoscope
 Genoscope - Centre National de Sequencage
 Bp 191 91006 EVRY cedex - France
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
 Location/Qualifiers
 1. 913
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_id="CS0DL012YD05"
 /clone_1lb="LTI_NFL010_BC2"
 /sex="male"
 /tissue_type="B cells from Burkitt lymphoma"
 /note="Vector: PCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the PCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

FEATURES

source

BASE COUNT 268 a 166 c 199 g 279 t 1 others

ORIGIN

Query Match 46.3%; Score 679; DB 106; Length 913;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 849; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

15 aaacccggagcgaggtctgctcttaaccgagcgccgctgctgctgagagacccgggtga 74
 |||||||
 1 AAACCCGGAGCGAGGTTCTGCTTACCCGAGCGCGCTGCTGCGGAGACCCCGGCTGA 60
 |||||||
 75 agcacccgtcatctgctcagcagagagcaaaacctcaactgaggaactgggggataa 134
 |||||||
 61 AGCACCGTATCATCTGCTGACCCAGGAGCAAAACCTTCAACTGAGCACTTGGGGGATTA 120
 |||||||
 135 gaagcaagtggaataataataaactcaaatcattggacagatagagagatcaact 194
 |||||||
 121 GAAGGAAGGTGAATATTTAACTCAAAAGTCAATTTGACAGAGATGAGTGAATTCATT 180
 |||||||
 195 caaagtgaataatgcaacacatctcaagaactcaagaactcaactgtaataaagaagg 254
 |||||||
 181 CAAGTGAATATGACACATCTCAAGAAATCAAGAAATCAATGATGATGATGATGATGAT 240
 |||||||
 255 tgtccaatgaatcactcaggtctcctttgagggagtcagagaatgctgataatacat 314
 |||||||
 241 TGTTCATGTAATTCATCTGAGGTTCTTTTGAAGGTCAGAGAAATGCTGATATATCATAC 300
 |||||||
 315 tccaaagaactgggaatggaagaagaatgtagatgagattatcaagaacaaacggg 374
 |||||||
 301 TCAGAAAGAACTGGGAATGAGAGAGAGATGTGATTTGAAAGTTATTCAGGAACAAACGG 360
 |||||||
 375 gggtcatcaacagttagatatattcttttctttcttcttccatcctttttt 434
 |||||||

```

Db 361 GGGTCATCAACAGTTAGATATCTTTTATTTTCTTTTCCCTCAATCCTTTT 420
QY 435 attttaaataagttcttctgtatgtgtgttcaaaagaaattgaacatgaccc 494
Db 421 ATTTTAAATATATCTTTTGTATGTGTTCACAAAGCAATGAAATGCGACACC 480
QY 495 catctcttgaacatctgtgaattgaattcagtcctcatatcatattgtt 554
Db 481 CATCTCTTTGAACATCTGCTAAATTTGAATTCAGTGTCTATTTATTTGTTGTT 540
QY 555 ttcatgttcgatttttggatgcaagccctcagtcctccctcatattacccctt 614
Db 541 TTCAATGTGCTGATTTTGGATCAAGCTTCAGTCCCTTCATATTACCTCTCTTT 600
QY 615 taaataatcgtgtgcaagaagagtcaccccttccaggaattgcatctcagctgt 674
Db 601 TAAATAATACGTGTGCAAGAGAGTGCACCTTTTCAGACATTCGATTTTCAGCTTGT 660
QY 675 ggtgataataagatcgaccaaagtgcaagtggtcataatgaattccaaatggt 734
Db 661 GGTGATTAATAGATCGACCAATGCACACTTTCATTAATGACTTTCACATTTGCCCTGATG 720
QY 735 ttc-agcatgtgattactactcctcctgacgtgacgtgacgtgagatggaattt 793
Db 721 TTTCAGCATGTGATTTACTTCTCTCTGACGTGTGACTTTTCAGTGGAGATGGAATTTT 780
QY 794 cagaagactgacgtgtgaaataatgaccccttcccttaactgaacttcaaatgt 853
Db 781 CAGAGAACTGAACTGTGAAATAATGACCTTTCCTTAACCTGAAGCTATTTAAATTTG 840
QY 854 agggctcgagc 864
Db 841 AGGGCTGAGC 851

RESULT 5
AL521484 868 bp mRNA EST 13-FEB-2001
LOCUS AL521484 LTI_NFL004_NBC2 Homo sapiens cDNA CSDB001YN17 5
DEFINITION prime, mRNA sequence.
ACCESSION AL521484
VERSION AL521484.1 GI:12784977
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 868)
AUTHORS L.M.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
1..868
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSDB001YN17"
/clone_id="LTI_NFL004_NBC2"
/sex="male"
/tissue="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain. Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(OT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by life technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center
Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : fliang@lifestech.com URL :
```

```

BASE COUNT 261 a 151 c 186 g 270 t
ORIGIN
Query Match 44.0% Score 645; DB 105; Length 868;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 865; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
QY 56 tgcgagaccccggtgaagcaccgcatcatgctcagcaggaaggaacacclta 115
Db 1 TCGGAGAGCCCCGGGTGAAGCAACCGTCATCATGTCTGACAGAGCAAAACCTTCA 60
QY 116 ctgagacttgaggagataaagaacaggtgatatataactcaactcaatggacag 175
Db 61 CTGAGACTTGGGGGATTAAGAAGAGGTGAATATTAATCAACTCAATCTGACAGG 120
QY 176 atagcagtgagattcactcaagaatgaaatgaaacacacatcctcaagaactcaagaat 235
Db 121 ATAGCGGTGAATTCATCTTCAAGTGAATGACACACATCTCAAGAACTCAAGAT 180
QY 236 catactgcaagaacaggtgttccaatgaaatcactcaagttctcttgaagggtcaga 295
Db 181 CATACTGCAAGAAGACAGGGTGTCCATGAATTCACCTCAGGTTCTCTTGAAGGTGAGA 240
QY 296 gaattgctgataataatactccaaagaactggaatggaagaagaatgtgttgaag 355
Db 241 GAATTTGCTGAATATCATCTCCAAAAGACTGGGAATGAGGAAGAAAGATGATTTGAG 300
QY 356 ttatcaagaacaaacgggggtcactcaacagtttaatalctcttatttcttct 415
Db 301 TTTATCAGGAACAAACGGGGGTCTTCAACAGTTTAAATTTATTTTATTTTCT 360
QY 416 ttccctcaatccttcttattttaaataagttccttggtaatggtgttcaaacg 475
Db 361 TTCCCTCAATCCTTTTATTTTAAATAAGTTCTTTGTAATGTGTCTCAAAAGC 420
QY 476 gaattgaaactgcaacccatcctcttgaacacatcgtgaatttgaattcagtgcca 535
Db 421 GAATTTAAACCTGGACCCCATCTCTTGAACATCTGTGTAATTTGAAATTTAGTCTCA 480
QY 536 ttatcattatgttcttcttcaatgtgtgatttggatgacagcctcagtcctcct 595
Db 481 TTATTCATTAATGTTTGTTCATATGCTGATTTTGTGATCAAGCTCAGTCCCTT 540
QY 596 catattaccctcctctttaaataatcagtgtagacagagaggtccactttttagac 655
Db 541 CATATTACCTCTCTCTTTTAAATAATGCTGTGCACAGAGAGGTCACTTTTTCAGGAC 600
QY 656 attgcatctcaagcttgtgtgataaataagatcgacaaatgcaagtgtcataatgac 715
Db 601 ATTGCATTTTCAAGGCTTGTGTGATTAATAAGATGACCAATGCAAGTTCATTAATGAC 660
QY 716 ttccaattggccctgattgttc-agcagtgaattacttcaactcctcgtgactgttca 774
Db 661 TTTCCAAATTTGGCCCTGAGATGTTCTACATGATTAATCTTCACTCCGAGCTGATTTTCA 720
QY 775 gtggagatggaagtttctcagaagaactgagtggaataatgaccccttcaactg 834
Db 721 GTGGAGATGGAAGTTTTCACAGAACTGAACGTGTGAAAAATGACTTTCCTTAACCTG 780
QY 835 aagctactttaaatttgaggtctggaacaaagaagaagatatcaagttgaagtca 894
Db 781 AAGCTACTTTTAAATTTGAGGGTCTGACCAAAAGAGAGAAATATCAGTTGAATGCA 840
QY 895 agatgacagataaggtgagatgac 922
Db 841 AGATGACAGATAAGCTGAGAGTAATGAC 868

RESULT 6
BG541610 835 bp mRNA EST 03-APR-2001
LOCUS BG541610
```


Db	Accession	Version	KeyWords	Source	Organism	Reference	Authors	Title	Journal	Comment
Db	149	AATATATTAACCTCAAGGTCATTGAGCAGATGATGACGATTCACCTTCAAAAGTGA	208							
Oy	206	tgaacacacatctcaagaacatcaagaatcatatctgtcaagaacaggtgttccaatga	265							
Db	209	TGACAAACACATCTCAAGAACTCAAGATATATCTGCTCAAGACAGGCTTCCAAATGA	268							
Oy	266	atcaactcaggtttctctctttgaggtcagagaattgtcgtatcatcatcactccaagaac	325							
Db	269	ATTCACTCAGGTTTCTCTTTGAGGCTCAGAAATTCGTGATATCATATCTCCAAAGAC	328							
Oy	326	tgggaatggaagaagaatgtagtltgaagittacagaaacaacggygggtcatcaa	385							
Db	329	TGGGATGTGAGGAGGAAGATGTGATTGAACTTATTCAGGAACAAACGGGGGTCAATCAA	388							
Oy	386	cagtttagatctctctttatctttttctttctttcccaatcccttttatttttaaaa	445							
Db	389	CAGTTTAAATATCTTTTATTTATTTTTCCTTCCTCAATCTTTTATTTTAAAAA	448							
Oy	446	tagtctcttctglaatgltgtgtlctaaaacggaatltgaaaacttgacacccactcttga	505							
Db	449	TAGTCTTTTGAATGTGTGTGTCAAAAACGGAATTTGAAACATGCGACCCCATCTTTGA	508							
Oy	506	aacatctgtaatttgaatcttcatgtctcatatcatcatattgtttgtttcattgtgc	565							
Db	509	AACATCTGTGTAATTTGAATTCATGTGCTCATTTATTCATTTATTTGTTTCATTTGCT	568							
Oy	566	gattttgtatcaagctcagtcacctcatatlaaccctctcctttttaaataacg	625							
Db	569	GATTTTGTGATCAAGCTCAGTCCCTTCATATTTACCTCCTCTTTTAAATATACG	628							
Oy	626	tgtcacagagaaggtcaccttttccagaacatgtcatlltcaggctgtgtgtataata	685							
Db	629	TGTGCACGAGAGGTGACCTTTTTCAGACATTCGATTTTCAGGCTGTGTGATAATA	688							
Oy	686	agatgaccacatgcagaatgttcataatgaacttccatltggccctgtatgtc	737							
Db	689	AGATCGACCAATGCAAGTGTTCATATATGATCTTCCAAITGGCCCTGATGTTC	740							
RESULT	8									
LOCUS	BE888192	858 bp	mRNA	EST	20-OCT-2000					
DEFINITION	601511710P1 NIH_MGC_71	Homo sapiens	cdna	clone IMAGE:3912914	5',					
ACCESSION	BE888192									
VERSION	BE888192.1	GI:10344249								
KEYWORDS	EST.									
SOURCE	human.									
ORGANISM	Homo sapiens									
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi									
AUTHORS	Mammalia; Euthelia; Primates; Catarrhini; Hominoidea; Homo.									
TITLE	1 (bases 1 to 858)									
JOURNAL	NIH-MGC http://mgc.nci.nih.gov/ .									
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)									
	Unpublished (1999)									
	Contact: Robert Strausberg, Ph.D.									
	Email: cgabbs-remail.nih.gov									
	Tissue Procurement: ATCC									
	CDNA Library Preparation: Life Technologies, Inc.									
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)									
	DNA Sequencing by: Incyte Genomics, Inc.									
	Clone distribution: MGC clone distribution information can be									
	found through the I.M.A.G.E. Consortium/LNL at:									
	http://image.llnl.gov									
	Plate: L14M9732	row: d	column: 03							
	High quality sequence stop: 726.									
FEATURES	location/Qualifiers									
source	1..858									

[illegible]


```

Db 422 AACCTTTTATTTTAAAAATAGTCTTTTGTAAATGATGTTCAAAACGAATTGAA 481
Qy 484 aacggaccccccctcttgaacacatctgttaattgaattcagaatcattatcat 543
Db 482 AACTGGACCCCACTCTTTGAACATCTGTAATTGTAATTCATGCTCATTAATTCAT 541
Qy 544 tatgttcttcttcatctgtctgtatttctgtgtacacagccctcagtcctcattac 603
Db 542 TATGTCGTTTCATGTCGTGATTTGGGTGATCAAGCTCAGTCCCTTCATATTAC 601
Qy 604 cctctcttctttaaataatcagtgacagagagtgacaccccttcttcagagcatt 663
Db 602 CCTCTCTTTTAAAAATTCGTGTGACAGAGGTCACTTTTCAAGCAATTCATT 661
Qy 664 ttcagctgtgtgtat 680
Db 662 TTGAGGCTGTGTGAT 678

RESULT 11
LOCUS AV718177 698 bp mRNA EST 16-OCT-2000
DEFINITION AV718177 FHFA Homo sapiens cDNA clone FHFAAH02 5', mRNA sequence.
ACCESSION AV718177
VERSION AV718177.1 GI:10815329
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 698)
AUTHORS Xiao,H., Peng,X., Song,H., Gu,Y., Yang,Y., Gao,G., Xu,X., Li,N.,
            Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L., Xu,S., Gu
            ,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Chen,Z. and Han
            ,Z.
            Homo sapiens cDNA FHFA clones
            Unpublished (2000)
            Contact: Zeguang Han
            Chinese National Human Genome Center at Shanghai
            351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
            201203, P. R. China
            Tel: 86-21-50801919(ex.45)
            Fax: 86-21-50801922
            Email: hanzg@hgc.sh.cn
            This clone is available at CHGC in Shanghai.
FEATURES
    source
        1. .698
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="FHFAAH02"
        /clone_1lb="FHFA"
        /tissue_type="hypothalamus"
        /dev_stage="Fetal"
        /lab_host="BM25.8"
        /note="Vector: pTRIPlex2; Site_1: sfiI; Site_2: sfiIbI"
BASE COUNT 201 a 127 c 155 g 213 t 2 others
ORIGIN
Query Match 34.1%; Score 500; DB 32; Length 698;
Best Local Similarity 99.8%; Pred. No. 3,7e-234;
Matches 550; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Db 128 ACTTGGGGGATTAAGAGCAAGGTGATATATTAATCAAGTCAATTGGACAGATAGCA 187
Qy 182 gtgagattcaactcaaaagtgaanaatgacaacacatctcaagaactcaagaatcact 241
Db 188 GTGAGATTCACTTAAAGTGAATGAAATGACACATCTCCAGAAACTCAAGAAATCAAT 247
Qy 242 gtcaaaacagaggtgtcccaatgaatcactcaggtctctcttgagggtcagaattg 301
Db 248 GTCAAGACAGAGGTGTTCCAAATGAAATTCACACTCAGGTTCTCTTGAGAGGTGAGAAAT 307
Qy 302 ctgataatcattaccacaaagaactggaaatggaggaagagatgtaattgaattatc 361
Db 308 CTGATTAATCAATCTCAAAAGAACTGGAAATGAGGAGGAAGATGATTAAGATTATC 367
Qy 362 aggaacaaacgggggtcattcaacagtttagatattcttcttcttcttcttcc 421
Db 368 AGGAACAACAGGGGGGTGATTCACACAGTTATGATATCTTTATTTTCTTTCC 427
Qy 422 tcaatccttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 481
Db 428 TCAATCTTTTATTTTAAAAATAGTCTTTGTATGTGTGTTCAAAACGGAATG 487
Qy 482 aaactggaccccccctcttcttgaacatctgtgaattgaattcagtgctcattatc 541
Db 488 AAAACTGGACCCCACTCTTTGAACATCTGTGAATTTGTAATTCATGCTCATTAATC 547
Qy 542 attatgtgt 552
Db 548 ATTAATGTTTG 558

RESULT 12
LOCUS BG399175 850 bp mRNA EST 12-MAR-2001
DEFINITION 602440659F1 NIH_MGC_75 Homo sapiens cDNA clone IMAGE:4556534 5',
            mRNA sequence.
ACCESSION BG399175
VERSION BG399175.1 GI:13292623
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 850)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
            Email: cgabs-remail.nih.gov
            Tissue Procurement: CLONTECH Laboratories, Inc.
            cDNA Library Preparation: CLONTECH Laboratories, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Cloning Distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNL at:
            http://image.llnl.gov
            Plate: LCM1259 row: e column: 15
            High quality sequence stop: 673.
FEATURES
    source
        1. .850
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="IMAGE:4556534"
        /clone_1lb="NIH_MGC_75"
        /lab_host="DH10B (TI phage-resistant)"
        /note="Organ: kidney; Vector: pDNR-LIB (Clontech); Site_1:
            SfiI (ggccgctggcc); Site_2: SfiI (ggccatattggcc); 5' and
            3' adaptors were used in cloning as follows: 5' adaptor
            sequence: 5'-CACGGCATTAATGCGC-3' and 3' adaptor sequence:
            5'-ATTCTAAGAGCCGAGGCGCCGAGATG-dr(30)BN-3' (where B = A,
            C, or G and N = A, C, G, or T). Average insert size 1.65
            kb (range 0.5-4.0 kb). 15/15 colonies contained inserts
            by PCR. This library was enriched for full-length clones

```


RESULT 14
LOCUS BG527341 822 bp mRNA EST 03-APR-2001
DEFINITION 602557324F1 NIH_MGC_59 Homo sapiens cDNA clone IMAGE:4686196 5',
mRNA sequence.
ACCESSION BG527341
VERSION BG527341.1 GI:13518878
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 822)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LNCM1496 row: p column: 05
High quality sequence stop: 800.
Location/Qualifiers
1..822
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4686196"
/clone_id="NIH_MGC_59"
/lssue_type="mucocutaneous carcinoma"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgctcgccg); Site_2: SfiI (ggcgctcgccg);
Double-stranded cDNA was prepared from cell line RNA. 5'
and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATTATGCGC-3' and 3'
adaptor sequence: 5'-ATTCTAGAGCGCGCGCGCGC-3' (30)BN-3'
(where B = A, C, G, or T). Average
insert size 1.65 kb (range 0.9-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
BASE COUNT 227 a 153 c 190 g 252 t
ORIGIN
Query Match 33.88; Score 496; DB 154; Length 822;
Best Local Similarity 99.88; Pred. No. 3.4e-232;
Matches 546; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Y 4 gaaagatttgaaccccgagaggttcgtctaccgagcgctgctgctgagaga 63
D 36 GAAGAGTTTGAACCCCGAGCGAGGTTCTTACCGAGCGCGCTGCTGCGAGGA 95
Y 64 ccccgagtgagacacgcgtcatcgtctgacagagagcaaaccttcaactgaagac 123
D 96 CCCCCGGGTAAAGCCACCGCATCATGTCTGACGAGGCAAAACCTTCAACTGAGAG 155
Y 124 ttgggggataaagcaagtgatgatatataactcaagtcattgagacagatagcagt 183
D 156 TTGGGGGATTAAGAGGATGATTAATTAATCAAAAGCATTTGACAGGATAGCAGT 215
Y 184 gaaattcaactcaaaatgaaatgacacacatcctcaagaactcaagaatcactgt 243
D 216 GAGATTCACTTCAAAAGTGAATGACAAACATCTCAAGAAATCAATCATACGT 275
Y 244 caaagacaggtgttccaatgaatcactcaagttctcttggaggtcagaagattgct 303

RESULT 15
LOCUS BG540101 804 bp mRNA EST 03-APR-2001
DEFINITION 602568937F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4693352 5',
mRNA sequence.
ACCESSION BG540101
VERSION BG540101.1 GI:13532334
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 804)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
CDNA Library Preparation: CLONTECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LNCM1515 row: j column: 09
High quality sequence stop: 785.
Location/Qualifiers
1..804
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4693352"
/clone_id="NIH_MGC_77"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: lung; Vector: pDNR-LIB (Clontech); Site_1:
SfiI (ggcgctcgccg); Site_2: SfiI (ggcgctcgccg); 5' and
3' adaptors were used in cloning as follows: 5' adaptor
sequence: 5'-CACGGCCATTATGCGC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGCGCGCGCGC-3' (30)BN-3' (where B = A,
C, G, or T). Average insert size 1.9
kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
by PCR. This library was enriched for full-length clones
and was constructed by Clontech Laboratories (Palo Alto,
CA). Note: this is a NIH_MGC Library."
BASE COUNT 227 a 150 c 184 g 243 t
ORIGIN
Query Match 33.88; Score 495; DB 155; Length 804;

Best Local Similarity 99.4%; Pred. No. 1,le-231;
Matches 695; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
OY 4 gaagatttgaacccccggagggaggttctgtaccgagggccgtgctgtgagaga 63
    |||||||
Db 19 GAAGATTGTGAACCCCGAGCGAGGTTGCTTACCCGAGCGCGTGTGTGGGAGA 78
OY 64 ccccggttgaagccaccgtcatcgtctgaccagaggaacacctcaactgagagac 123
    |||||||
Db 79 CCCCCGGTGAAGCCGTCATCATGTCTGACCAAGAGGCAAAACCTTCAACTGAGGAC 138
OY 124 ttggggtaagaagcaagtgaaataataaactcaaatgcaatggacagagagc 183
    |||||||
Db 139 TTGGGGGTAAGAAGAGAGTGAATATATTAACCTCAAGTCATTGGACAGATAGCAGT 198
OY 184 gagatccttcaagaatgaataatgacaacacatctcaagaactcaagaatcactgt 243
    |||||||
Db 199 GAGATTCACTCAAAAGTAAGTAACACACACATCTCAAGAACTCAAGATCACTGT 258
OY 244 caaagacagaggtgttccaatgaatcaactcaggttctcttggaggtcagagatgct 303
    |||||||
Db 259 CAAGACAGAGGTTGCCATGATTCATCTGAGGTTCTCTTGGAGGTCAGAGATTGCT 318
OY 304 gataatcactccaagaactgaggaaatggaagaagaatgtgattgaagttatcag 363
    |||||||
Db 319 GATATATCTACTCCAAAAGAACTGGGAATGAGAGAGAAATGTGATTGAAGTTATCAG 378
OY 364 gaacaacggggggggttcaactcaagttagatcttatttttttcttccctc 423
    |||||||
Db 379 GAACAAACGGGGGTCATTCAACAGTTAGATATCTTTTATTTTCTTTCCCTC 438
OY 424 aatccttttattttaaataatagttccttctgtaalggtgttcaaaacggaattgaa 483
    |||||||
Db 439 AATCCTTTTATTATTTTAAATAATAGTCTTTGTAATGTGTTCAAAACGGAATTGAA 498
OY 484 aactggacccccactctcttgaacaatctggttaattgaaatctagtgctcatatcat 543
    |||||||
Db 499 AACTGGCACCCCATCTCTTGAACATCTGTGAATTGAAATCTAGTGCATATTTCAT 558
OY 544 tatgttgttcatatgctgctgatttttggtagtcaagcctcagtcctccatattac 603
    |||||||
Db 559 TATGTTTGTTCATTTGCTGTGATCCTTGTGATCAAGCCTCAGTCCCTTCATATTAC 618
OY 604 cctctcctttaaataatlaacgtgtgacaagaaggtcaaccttttcaagacattgcat 663
    |||||||
Db 619 CCTCTCCTTTTAAATAATAGTGTGACAGAGAGGTACACTTTTCAGAGCATTTGCATT 678
OY 664 ttcaaggtctgtgtgataaataaagatcgaccaatgcaag 702
    |||||||
Db 679 TTCAGGCTTGTGTGATTAATTAAGATGACCAATGCAAG 717
```

Search completed: August 17, 2001, 10:45:41
Job time: 4517 sec

